GreatTurbo Cluster Server 10

用户快速安装手册



目录

介绍

注意事项

<u>在您开始使用 GreatTurbo Cluster Server 10 之前，请您全面仔细的阅读本用户手册！否则，可能会因为您的疏忽而使得安装过程出现不必要的问题！</u>

GreatTurbo Cluster Server 10 介绍

GreatTurbo Cluster Server 10 是北京拓林思软件有限公司推出的为满足 Linux 平台企业级应用的高可用和负载均衡产品。它同时具备了高可用和负载均衡产品所有的功能，而且在原有功能的基础上增加了许多新的功能。它提供的双机高可用方案在保证负载均衡调度器高可用的同时，还能够满足其他应用对高可用的要求。其良好的可靠性、可用性、可扩展性为用户业务提供了强有力的支持和保证。

1. GreatTurbo Cluster Server 10 的使用限制

- ✧ GreatTurbo Cluster Server 10 目前只支持 GTES10/TDS10 操作系统平台。
- ✧ GreatTurbo Cluster Server 10 暂不支持并行处理的应用。也就是说不支持同一个应用在两个节点同时并发运行的应用。例如：Oracle 9i RAC。
- ✧ GreatTurbo Cluster Server 10 配置的所有负载均衡(LB)服务都必须采用同一种 ip 负载均衡技术。
- ✧ GreatTurbo Cluster Server 10 配置的所有 LB 服务都必须运行在同一调度节点，并且只能对全部服务进行统一操作，如 Enable、Disable、Relocate 等。
- ✧ GreatTurbo Cluster Server 10 的稳定性需要 OS 提供支撑。如果当操作系统宕机时，可能会出现因 OS 没有彻底释放资源而导致 GreatTurbo Cluster Server 10 系统丧失高可用功能。在这种情况下，除非有额外的硬件设备，否则 GreatTurbo Cluster Server 10 并不能够完全保证能够自动恢复操作系统。此时需要用户手工干预操作系统，对崩溃的操作系统进行复位操作。也就是说，当 OS 宕机时（尽管这种可能性很小），如果用户没有以下硬件作为保障，仍然有可能会出现用户业务中断的可能：

  1）采用电子开关。

  2）服务器节点采用支持硬件 watchdog 功能的主板。
- ✧ 配置 GreatTurbo Cluster Server 10 的两台计算机节点的心跳方式时，必须保证至少一条心跳通道正常工作。如果两台节点之间的所有心跳通道都发

生故障而不能正常连通，有可能会导致 GreatTurbo Cluster Server 10 发生裂脑（split-brain），发生裂脑后，GreatTurbo Cluster Server 10 有可能会导致用户的资源不一致。为了完全杜绝裂脑现象的发生，可以采取如下方法：

1） 采用电子开关。

2） 使用第三方参考 IP，有关第三方参考 IP 的介绍将在第三章详述。

其中第一种办法是使用硬件的办法,由于电子开关是额外的电子硬件设备，需要用户自行购买，所以实际中采用这种方式并不多；而第二种方式是软件的方法，可以保证 GreatTurbo Cluster Server 10 发生裂脑时，用户的资源不受损失。这需要用户提供一个能长期(永久性)正常工作的参考 IP 地址。

2. GreatTurbo Cluster Server 10 的相关术语

节点 ： 指运行 GreatTurbo Cluster Server 10 程序的计算机。

服务 ： 也叫做**资源组**，指用户应用相关的一组资源的集合，包括用户应用的进程资源，磁盘资源，网卡资源，浮动 IP 资源，drbd 镜像资源等。服务可以是其中几种资源或者全部资源的组合。服务也可以为空，即不包括任何资源。通常用户的一个应用与 GreatTurbo Cluster Server 10 的一个服务对应。GreatTurbo Cluster Server 10 的服务分为两种：高可用服务（HA 服务）和负载均衡服务（LB 服务）。HA 服务针对高可用集群确保高可用的对象。LB 服务是负载均衡系统进行调度的具体应用，例如 Oracle 服务。HA 服务也可以是具体商业应用。

负载调度器（load balancer）：也成为调度节点，它是整个负载均衡集群对外的前端机。负责将客户的请求发送到一组真实服务器上执行，而客户认为服务是来自一个 IP 地址(我们可称之为虚拟 IP 地址)上的。在 GreatTurbo Cluster Server 10 典型应用中，负载调度器需要 2 个或多个，GreatTurbo Cluster Server 10 的高可用方案保证负载调度器能稳定、24×7 不间断的运行。

真实服务器 （real server）：也叫服务器池（server pool），是负载均衡集群中真正执行客户请求的服务器，执行的服务有 WEB、MAIL、FTP 和 DNS 等。用户请求由调度节点分发给真实服务器，由真实服务器来处理用户请求。

主节点 ： 指服务运行所在的节点。

备节点 ： 指完全没有服务运行的节点。如果主节点发生任何故障，服务就会从主节点迁移到备节点。此时的备节点也就转变成主节点，此前的主节点也就转变成备节点。

主备方式 ： 常见的主备方式有 Active-Standby、Active-Active。Active-Standby 是指服务仅在一个节点上运行。Active-Active 是指在两个节点都有服务运行。也就是说，Active-Active 方式时都是主节点，也都是备节点，互为主备关系。

优先节点（preferred node）： 指服务将优先运行的节点。当配置一个服务时，可以给这个服务设定优先节点。一般有两种情况建议配置优先节点，一种情况是当两个服务器节点的硬件配置不一样时，应当将服务的优先节点设定为硬件配置较好的节点；另一种情况是当需要配置多个服务且应用的负载较重，而且希望两个

服务分别在两个节点上运行以分担负载时，可以给这两个服务分别各设定一个节点作为优先节点。

例如：假设服务 s 配置的优先节点是节点 A，并且将服务 s 在 A 节点启动，当 A 节点发生故障时，服务 s 会迁移到 B 节点，后来当 A 节点恢复之后，服务 s 会因为 A 节点设置成优先节点而自动回迁移到 A 节点。

当然，配置优先节点带来的弊病是，会导致服务多迁移一次。

服务的迁移（relocate）： 是指服务在一个节点发生故障之后，服务先在故障节点进行服务的停止以释放服务的所有资源，然后在另一节点启动服务，使服务继续可用的过程。

服务迁移的时间 ：是指服务不可用的时间，也就是服务正在恢复之中但尚不可用的时间段。

例如：服务在 A 节点发生故障的时刻为 T(a)，服务自动迁移到 B 节点并在 B 节点成功运行的时刻为 T(b),那么服务的迁移时间 T(relocate)=T(b)-T(a)。另一方面，从整个迁移的子过程来看，服务迁移的时间（近似等于）≈服务在 A 节点检测到错误的时间＋服务在 A 节点停止的时间＋服务在 B 节点启动的时间。其中，服务的启动/停止时间是由用户的应用来决定的，一般不能够进行调整，而服务检测到错误的时间可以通过 GreatTurbo Cluster Server 10 的配置参数来进行调整。

服务检测到错误的时间 ＝ 服务检查的间隔时间（check interval） * 服务连续检查到错误的次数（check count）。Check interval 和 check count 参数可以在配置服务的检测脚本时进行指定。这样，只有当 GreatTurbo Cluster Server 10 检测到连续出错次数达到指定的次数后，GreatTurbo Cluster Server 10 才认为服务确实出错。服务确实出错后，会自动触发 GreatTurbo Cluster Server 10 进行服务的迁移。

裂脑（split-brain）： 所谓裂脑，是指 GreatTurbo Cluster Server 10 的高可用主备节点之间彼此失去了联系,但是单个节点的 GreatTurbo Cluster Server 10 仍然运行正常。发生裂脑的充要条件是：1. 系统的主备节点之间的所有心跳通道都发生了故障,导致系统的两个节点失去了任何联系。2. 两个节点的 GreatTurbo Cluster Server 10 软件仍在正常运行。

裂脑带来的直接后果是导致两个节点都会各自启动并运行同一个服务，竞争同一个服务的资源，这样有可能会导致资源（尤其是共享数据）被损坏。

drbd（磁盘镜像设备） ： 是指对两个节点各自大小相同的本地磁盘分区进行实时镜像操作，使得这两个本地磁盘对双方节点而言，可以当作一个虚拟的共享磁盘阵列来使用，这个虚拟的共享磁盘阵列就叫做 drbd 设备，可以把 drbd 当作普通的磁盘设备来使用。

Watchdog（看门狗） ： Watchdog 分为硬件级和软件级两种。硬件级 watchdog 是用来保障操作系统自动恢复的一种手段，如果服务器节点的主板支持 watchdog 功能，那么在 GreatTurbo Cluster Server 10 中可以进行相应的配置，当操作系统发生死机等不响应情况，主板就会将操作系统自动重启恢复，而无需人工干预，这一点对于“24 * 7”方式运行的服务器而言非常有用。而软件级的 watchdog 可以用来保证 GreatTurbo Cluster Server 10 程序的健壮性但并不能保障操作系统的自动重启恢复。所以硬件级的 watchdog 更加实用。

3. GreatTurbo Cluster Server 10 快速安装手册的使用方法

由于本手册是快速安装文档，所以各个章节的部分基本上都需要仔细阅读。其中第五章至第八章，第九章至第十二章根据您的应用类型选择进行阅读即可，其余章节都是必读部分。

4. GreatTurbo Cluster Server 10 快速安装的主体过程

GreatTurbo Cluster Server 10 快速安装的主体过程有如下六个步骤：

1) 进行安装/使用 GreatTurbo Cluster Server 10 前的准备工作。（第一章介绍）
2) 安装 GreatTurbo Cluster Server 10 软件。（第二章介绍）
3) 配置 GreatTurbo Cluster Server 10 软件。（第三章介绍）
4) 准备服务的用户脚本和检测脚本。（第四章介绍）
5) 根据您的应用类型，将应用配置成高可用的服务。（第五/六/七/八章介绍）
6) 根据您的应用类型，配置具体的 LB 服务（第九/十/十一/十二章介绍）
7) 监控和维护 GreatTurbo Cluster Server 10 系统。（第十三、十四章介绍）

另外本手册在第十五章介绍了如何卸载 GreatTurbo Cluster Server 10，并且在附录中还列出了 GreatTurbo Cluster Server 10 安装使用过程中常见的问题以及如何进行一些后期的维护。

第 1 章

安装 GreatTurbo Cluster Server 10 前的准备工作

安装 GreatTurbo Cluster Server 10 之前需要您做好相关的软硬件环境的准备工作。否则，可能会影响您随后的安装过程，使您遇到不必要的麻烦。

安装前的准备工作可以分为如下四个步骤：

1) 了解和分析您的应用；
2) 准备 GreatTurbo Cluster Server 10 需要的硬件环境；
3) 准备 GreatTurbo Cluster Server 10 需要的软件环境；
4) 安装 GreatTurbo Cluster Server 10 前需要准备的操作系统的配置。

1. 了解和分析您的应用

GreatTurbo Cluster Server 10 可以支持绝大多数的双机高可用和负载均衡的应用。按照服务类型，GreatTurbo Cluster Server 10 将用户的应用分为两大类：HA 应用和 LB 应用。

1）HA 应用

根据用户应用是否需要使用共享的数据资源，GreatTurbo Cluster Server 10 将用 HA 应用分为以下三类：

a)第一类应用：需要使用共享磁盘阵列的应用。

这类应用的特点是，应用的数据非常重要且 I/O 负载较高，必须使用共享磁盘阵列来保证应用数据的一致性、可靠性和高性能。此类应用的典型例子是各类数据库应用，如 Oracle、Sybase、MySQL 等。也就是说，使用 GreatTurbo Cluster Server 10 的双机系统本身作为后台数据系统。

b)第二类应用：不需要使用共享磁盘阵列并且也不需要存取共享数据的应用。

这类应用的特点是,不需要存取共享数据,使用 GreatTurbo Cluster Server 10 的双机系统只是用于进程和浮动 IP 等资源的切换。一般情况下，这种类型的应用只是用来作为前端接入、前置机来使用。典型的应用有 WEB/MAIL 等网络应用或者是用户自己的业务应用。这类应用通常会有另外的后台数据库系统来保存数据。

c）第三类应用：需要使用共享磁盘阵列但是不具备磁盘阵列硬件条件的应用。

这类应用的特点是 I/O 负载较小（<10MB/s），数据本身并不是关键业务数据，并且该应用需要使用共享的存储设备来保证数据的一致性。对于此类应用，GreatTurbo Cluster Server 10 提供了虚拟的共享磁盘设备“磁盘镜像设备”来保证应用数据的一致性。

2）LB 应用

根据使用的负载均衡技术，LB 应用分为以下三类：

a)第一类应用：使用 VS/DR 负载均衡技术的应用。

b)第二类应用：使用 VS/TUN 负载均衡技术的应用。

c)第三类应用：使用 VS/NAT 负载均衡技术的应用。

注意：

通常情况下，用户的一个应用对应于 GreatTurbo Cluster Server 10 的一个服务。如果有多个应用就需要在 GreatTurbo Cluster Server 10 中配置多个服务，依照每个应用的类型，您可以分别按照第五、六、七、八章，第九、十、十一、十二章的内容进行配置。

2．准备 GreatTurbo Cluster Server 10 需要的硬件环境

1） 两台 X86 架构的计算机，硬件的配置只需满足能够安装主流的 Linux 操作系统的条件即可。推荐用户使用具备 RAID-1 级别本地磁盘设备的服务器。

2） 虽然 GreatTurbo Cluster Server 10 本身对硬件的要求不高，但是节点的硬件配置需要和用户的应用负载相匹配，也就是说用户的硬件环境的性能配置应该满足用户的应用负载的峰值要求！

3） 每台计算机必须配备三块以上网卡。其中，用作两节点心跳信息的网卡采用直连网线相连接，至少需要两条直连网线作为心跳线。

网卡的性能要求取决于您的应用的类别。如上所述，如果您的应用类别是第三类应用时，建议您至少需要使用一块 1Gbps 的网卡作为直连网线的设备，这样可以提高磁盘镜像设备的性能；如果是第一类应用或者是第二类应用时，网卡性能达到 100Mbps 即可（当然如果您的应用对网络带宽要求较高，则也需使用 1Gbps 网卡）。

请将两台计算机的所有用作 GreatTurbo Cluster Server 10 通道的对应网卡用直连网线一一连接起来。

4） 一根串口线。为了保证系统的高可用性，需要再使用一根串口线连接两个节点的串口。

请将两台计算机对应的串口用串口线连接起来。

5） 是否需要共享的磁盘阵列，取决于您的应用。如果您的应用是第二类或第三类 HA 应用，您可以不需要使用磁盘阵列。如果您的应用是第一类应用，您需要

使用共享的磁盘阵列。另外，如果您需要使用共享 raw 磁盘分区作为心跳方式，则需要使用磁盘阵列

在选用磁盘阵列时，一般分两种情况：第一种是带独立 RAID 处理器的磁盘阵列，主流厂商的 SCSI 或光纤磁盘阵列都可以满足 GreatTurbo Cluster Server 10 的要求，这类磁盘阵列产品的性能强且价格也较高。第二种是使用主机 RAID 卡和磁盘柜（磁盘柜是指不具备硬件 RAID 处理器的磁盘盒）的方式，这种磁盘设备不是真正意义上的磁盘阵列，但通过双机的 RAID 卡 clustering 技术和 RAID 的用户接口工具等，也可以满足 GreatTurbo Cluster Server 10 的共享数据存取的需求，常见的这种类型的“磁盘阵列”的典型产品有 IBM 的 EXP400、DELL 的 220S、HP 的 MSA500-G2 等，这类“磁盘阵列”产品的性能和价格一般都较前一种低。

注意：

1）在配置磁盘阵列和做 RAID 盘的时候，请认真仔细参考磁盘阵列提供厂商的手册，根据应用的特点来配置相关参数，如 hotspare 盘、raid 级别、stripe 大小、cache 策略等。如果不进行配置参数的调优将会导致磁盘阵列的可靠性和 I/O 性能受到较大影响。

2）在使用第二种“磁盘阵列”的时候，一定要仔细询问相应的硬件厂家，其产品是否支持双机 GreatTurbo Cluster Server 10 的数据共享应用。另外，建议尽量选择那些缓存较大同时附带后备电池的 RAID 卡型号，以提高 RAID 卡的高可用性。

3）无论使用那种磁盘阵列，都一定要和硬件厂商确认，该磁盘阵列相关设备的驱动程序是否为最新的稳定版本，以提高磁盘阵列的高可用性。

3. 准备 GreatTurbo Cluster Server 10 需要的软件环境

安装 GreatTurbo Cluster Server 10 的软件条件比较简单，必需要安装 GTES10 操作系统平台。

如果您想使用图形界面的管理工具，那么必须保证在客户端和服务器端的操作系统上提供 Java 运行时环境支持，JRE1.4.2 及其以上版本。如果您不具备 java 环境，那么也可以使用文本方式的管理工具进行配置。

注意：

用户在安装操作系统时，建议用户采用完全安装操作系统的模式，并且将操作系统安装在本地具有 RAID-1 功能的磁盘设备上。这样可以保证操作系统功能的完备性和操作系统数据的可靠性。

如果使用图形管理工具，检查/usr/bin 目录下的 java 是否正确链接到 JRE 目录。

4. 安装 GreatTurbo Cluster Server 10 前需要准备的操作系统的配置

两个节点的操作系统配置需要确认如下环节：

1) 请确认您登陆操作系统的身份是 root 权限，因为随后的安装配置都需要是 root 权限。
2) 如果您需要使用共享的磁盘阵列，请查看磁盘阵列是否已经可以使用（包括硬件的配置、RAID 的配置、驱动的安装、分区、文件系统等）。
3) 所有网卡是否配置好。（包括直连网线是否连好、hosts 文件是否设置等）。如果使用了串口，查看双方的串口是否正常连接。
4) 由于 GreatTurbo Cluster Server 10 使用了操作系统的 syslog 服务来打印 GreatTurbo Cluster Server 10 的 LOG 日志，所以请确认 syslog 服务已经配置并且正在运行。例如：

```
[root@test1 root]# /etc/init.d/syslog status
syslogd (pid 447) is running...
klogd (pid 457) is running...
```

5) 如果你需要使用共享磁盘阵列,并且需要在共享磁盘上 make 文件系统，除非特殊的要求,强烈建议文件系统的类型为日志文件系统(例如 EXT3 等）。同时 make 文件系统时，注意调整 block 参数的大小以便与磁盘阵列做 RAID 盘时的 stripe 参数对应成最佳。这样可以提高文件系统的性能。
6) GreatTurbo Cluster Server 10 的日志系统要求两台节点的系统时钟一致，以方便后期维护。如果两台系统时钟不一致，请使用 hwclock 以及 date 命令来调整时钟，使得两台节点的时钟大致相同，误差小于 30 秒之内。
7) 为了方便后期维护，需要系统的 sshd 服务正常运行，并且允许 root 权限进行 ssh 登陆。

5. 安装 GreatTurbo Cluster Server 10 前需要准备用户的应用程序

接下来，请在两个节点上安装用户的应用程序，两个节点的安装配置都要求是对等的，也就是说，<u>在两个节点上安装的用户的应用程序所在的目录、使用的共享磁盘设备名称、网卡名称等都应该一致。</u>

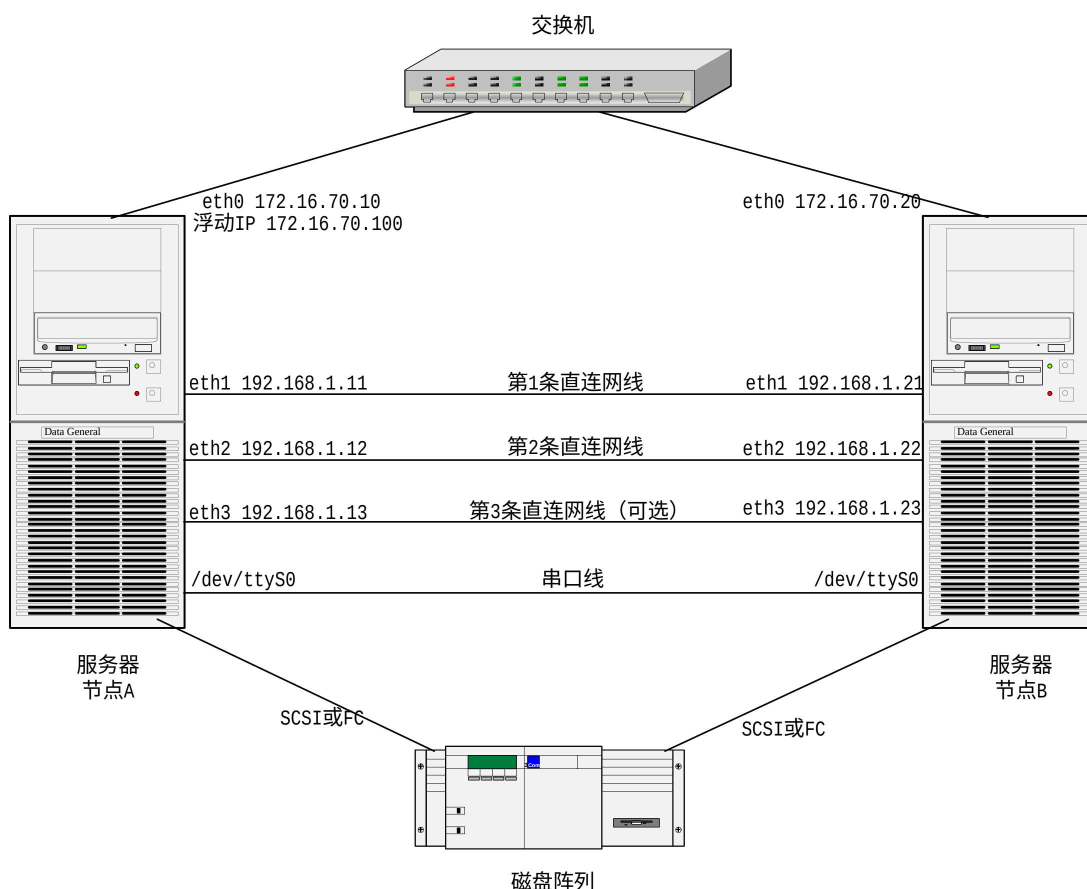
通常，都会将应用程序安装在两个节点的本地硬盘上。

如果如果您的应用类别是第一类应用时，可能还需要在共享磁盘设备上准备共享的数据。例如：对于 oracle 应用，您需要表空间文件存放在共享磁盘设备上。

上面介绍了安装前需要的准备工作，接下来将通过具体的例子来说明。

6. 安装 GreatTurbo Cluster Server 10 的配置示例

现在我们引入一个较为通用的双机软硬件环境的例子来说明具体的过程。在随后几章的安装、配置、使用等章节中的具体内容将参照如下拓扑结构的情况进行说明。



如上图所示，有两台节点，服务器节点 A 和节点 B。假设每台机器有四个网卡 eth0-eth3，eth0 连接至交换机，eth1-eth3 和另一节点的对应网卡通过直连网线相连，并且两个节点通过串口线将各自的第一个串口相连。此外，两台机器通过 SCSI 线缆与共享的磁盘阵列相连。

在节点 A/B 两个节点上分别配置好各个网卡的信息（包括 IP 地址、网络掩码、广播地址等），分别将 eth0 的 hostname 设置成主机的 hostname，假设节点 A 的 hostname 对应于 eth0，且设置成 test1，节点 B 的 hostname 对应于 eth0，且设置成 test2。并且分别在/etc/hosts 文件中添加两个主机 eth0-eth3 网卡的别名。例如 hosts 文件内容如下：

```
172.16.70.10            test1
172.16.70.20            test2

192.168.1.11            hb11
```

```
192.168.1.12        hb12
192.168.1.13        hb13
192.168.1.21        hb21
192.168.1.22        hb22
192.168.1.23        hb23
```

注意：不要把 hostname 对应到 127.0.0.1 这个 IP 地址。

其中 hb11、hb12、hb13 对应节点 A 的 eth1、eth2、eth3 三个网卡；hb21、hb22、hb23 对应节点 B 的 eth1、eth2、eth3 三个网卡。

第 2 章

安装 GreatTurbo Cluster Server 10

在做好了第一章中所列的各项准备工作之后，现在就可以安装 GreatTurbo Cluster Server 10 软件了。

安装过程如下：

1) 确认您所使用的 GreatTurbo Cluster Server 10 产品的功能级别
2) 安装 GreatTurbo Cluster Server 10
3) 注册 GreatTurbo Cluster Server 10

1. 确认您所使用的 GreatTurbo Cluster Server 10 产品的类型

GreatTurbo Cluster Server 10 产品分为两个功能级别，请确认您购买的产品属于哪一个级别。第一级别只能够满足第一章提到的 HA 应用的第一类和第二类的用户应用，第二级别能够全部满足第一章提到的三类 HA 应用。也就是说，第一级别不支持磁盘镜像设备（drbd）功能，第二级别支持 drbd 功能。

当然，具体选购何种级别功能的 GreatTurbo Cluster Server 10，完全取决于您需要满足的应用类型。确认了 GreatTurbo Cluster Server 10 的功能级别之后，就可以开始安装 GreatTurbo Cluster Server 10 了。

2. 安装 GreatTurbo Cluster Server 10

请分别在节点 A 和节点 B 安装 GreatTurbo Cluster Server 10 软件。

1) 如果您购买的是正版的 GreatTurbo Cluster Server 10 产品，GreatTurbo Cluster Server 10 会附带软件的安装光盘。插入安装光盘到节点的光驱并 mount 后，光盘根目录中有一个安装文件 install_cluster，请运行如下命令进行安装。根据您的操作系统版本并参照其提示完成 GreatTurbo Cluster Server 10 的安装。

如果您的 GreatTurbo Cluster Server 10 产品支持 drbd，请在“Do you want to use drbd ?”时选择 yes，否则选择 no。

例如：

```
[root@test1 root]# mount  /dev/cdrom  /mnt/cdrom
[root@test1 root]# cd  /mnt/cdrom
[root@test1 cdrom]# ./install_cluster
Following RPMs will be installed or upgraded to newer version if
necessary:
```

```
  *) pdksh
  *) lsof
  *) sg_utils
  *) drbd
  *) GreatTurbo Cluster Server
  *) GreatTurbo Cluster Server Agents
Please select the operatating system version you are using:
  0) GreatTurbo Enterprise Server 10 (x86_32)
  1) GreatTurbo Enterprise Server 10 (x86_64)
  2) GreatTurbo Enterprise Server 10 (OpenPower ppc64)
  3) GreatTurbo Enterprise Server 10 (IA64)
  4) cancel
Please select an OS version, then the installation procedure will begin.
Select 4 to abort [0/1/2/3/4]: 0
Do you want to use drbd? (y/n) [y]: y
Please select the network forwarding method you want to use:
  0) direct routing
  1) NAT
  2) IP tunnel
Please select a network forwarding method [0/1/2]: 1
```

2） 如果您是从拓林思公司下载的试用产品（使用有效期为 1 个月），您下载的 GreatTurbo Cluster Server 10 将是一个 ISO 文件和 md5 检验文件。ISO 文件的名称一般为 greatturbocluster-10.x-x.iso，md5 文件的名称一般为 greatturbocluster-10.x-x.iso.md5。为了验证您下载的 iso 文件是否正确，可以先对您的 iso 文件进行 md5sum 操作，如果发现 md5sum 运行命令所得结果，和您下载的 md5 文件的内容一致，则表明您下载的 ISO 文件正确无误。

例如：

```
[root@test1 iso]#md5xum ./greatturbocluster-10.0-1.iso
88f7b387c8cdaca3fa38832c6b7b2b6a  greatturbocluster-10.0-1.iso
[root@test1 iso]# cat ./ greatturbocluster-10.0-1.iso.md5
88f7b387c8cdaca3fa38832c6b7b2b6a  greatturbocluster-10.0-1.iso
```

证明您下载的 ISO 文件正确之后，请先将 ISO 文件 mount 到某一个目录，然后运行这个目录下的 install_cluster 安装程序进行安装，接下来的过程同从光盘安装时介绍的一样。

```
[root@test1 root]# mount –o loop  ./ greatturbocluster-10.0-1.iso
/mnt/iso
[root@test1 root]# cd  /mnt/iso
[root@test1 iso]# ./install_cluster
```

3. 注册 GreatTurbo Cluster Server 10 产品

安装完 GreatTurbo Cluster Server 10 的软件之后,您需要注册 GreatTurbo Cluster Server 10 产品，也就是说需要安装 GreatTurbo Cluster Server 10 的 license，以保证 GreatTurbo Cluster Server 10 正常运行。

注册 GreatTurbo Cluster Server 10 产品的步骤如下：

1) 获得两个节点的硬件号,分别在两个节点上运行”/opt/cluster/bin/syncd -l” ，其输出的第一行信息，例如“Hardware ID: 3355DEJGQWVK”，其中“3355DEJGQWVK”就是硬件号。仔细记录两台节点的硬件号，以便在随后的注册中使用。
2) 用浏览器登陆注册网站 http://www.turbolinux.com.cn/register。进入如下图所示的登陆页面：
3) 如果您是第一次注册，请点击“新用户”按钮，按要求详细填写注册信息。
4) 注册完之后，返回如上登陆页，输入电子邮件地址和密码，点击“登陆”，进入注册页。您还可以在产品序列号栏，输入您购买产品时附带的序列号，以查看序列号有没有被注册过，如果发现被人盗用注册过，则请与拓林思软件有限公司联系，因为注册码只允许注册一次。
5) 登陆后，进入以下页面：

如果您是正版用户，请点击“增加产品序列号”按钮、进入步骤 6）。如果您没有序列号，只是试用，请点击“评估版许可证”按钮、进入步骤 7）。

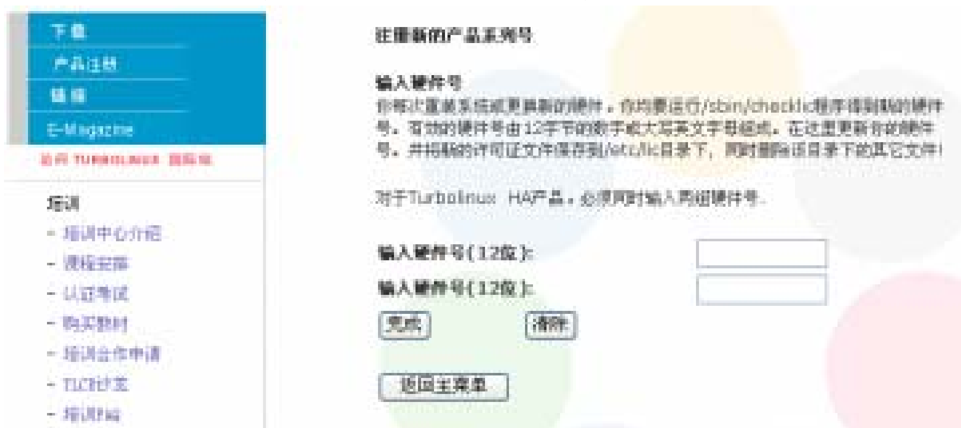
6） 点击“增加产品序列号”之后，进入如下页面：



7） 在 SN 栏输入您购买产品时附带在注册卡上的序列号、然后点击“下一步”按钮，或者点击 5）中“评估版许可证”进入如下页面：



8） 选择您的产品类型，点击“下一步”按钮，进入以下页面：



9） 分别填入两个节点的硬件号，然后点击完成，进入到如下所示的页面：



10) 点击“ 取得使用许可证 ”按钮，在接下来的页面上选择如何保存你得许可证：寄到信箱或者存到文件。
11) 得到 license 文件之后，你需要把它分别 copy 到两个节点的/etc/opt/cluster/lic 目录下。

注意：

在/etc/opt/cluster/lic 目录下只能保存一个 license 文件。

至此，GreatTurbo Cluster Server 10 产品注册完毕。接下来在第三章中，我们将讲述如何对 GreatTurbo Cluster Server 10 进行初始配置。

第 3 章

系统配置

在以上步骤配置完成之后，就可以初始化 GreatTurbo Cluster Server 10 了。初始化过程是通过 **member_config** 命令来完成的。

初始化 GreatTurbo Cluster Server 10 的过程如下：

1）选择其中一节点（例如节点 A）进行系统配置，

2）在另一节点（例如节点 B）同步配置信息。

注意：

member_config 只需要在一个节点上配置；系统运行中不能使用 member_config。

1. 选择其中一节点进行系统配置

首先我们需要选择其中一节点对 GreatTurbo Cluster Server 10 进行一些配置。

1）启动 member_config 命令

注意：

如果您使用了 IBM 的 EXP400 系列的磁盘柜，在确定 IBM 的 EXP400 磁盘阵列的硬件安装和配置正确之后，运行 member_config -s 命令。

运行 member_config 命令，系统显示如下：

```
[root@test1 root]# /opt/cluster/bin/member_config
------------------------------------
Cluster Member Configuration Utility
Version：10.0 Built: Fri Jun 9 16:51:41 CST 2006
This utility sets up the member systems of a 2-node GreatTurbo Cluster
Server 10 cluster,
or the 2-node director members of a Load Balancing cluster.
It prompts you for the following information:

o  Hostname
o  Number of heartbeat channels
```

```
o  Information about the type of channels and their names
o  Power switch type and device name
o  Information about the routers and network type of the Load Balancer

In addition, it performs checks to make sure that the information
entered is consistent with the hardware, the Ethernet ports, the raw
partitions and the character device files.

After the information is entered, it initializes the configure file
and saves the configuration information to the configure file

- Checking that cluster daemons are stopped: done
- Configuration file exists already.
  Would you like to use those prior settings as defaults? (yes/no) [yes]:
```

如果以前运行过 **member_config** 命令，那么配置文件中会保存我们的配置结果，该选项就是问我们是否使用以前的配置结果作为缺省值，通常回答 yes。

注意：如果是第一次配置，则不会出现该选项。

2）配置本地节点名称

接下来，输入本地节点的名称。GreatTurbo Cluster Server 10 会自动的从您系统中/etc/hosts 中得到本地节点的名称，前提是您没有把这个名称对应到 127.0.0.1。 如果这个名称对应到 127.0.0.1，安装程序会退出，并提醒您正确的配置/etc/hosts。

```
----------------------------------------
Setting information for cluster member 0
----------------------------------------
Enter name of cluster member [test1]: test1
Looking for host test1 (may take a few seconds)...
Host test1 found
Cluster member name set to: test1
```

3）配置本地节点的 heartbeat 选项：

GreatTurbo Cluster Server 10 的 heartbeat 通道有两类： 网络(net)和串口(serial)。对于网络，您需要配置 heartbeat 使用的网络设备对应的别名；对于串口，您需要配置 heartbeat 使用的串口的设备名，例如 /dev/ttyS0。

注意：为了获得更高的可用性，GreatTurbo Cluster Server 10 建议配置一条串口通道以及至少两条直连网络心跳通道，并且必须将应用所在的网卡配置成通

道，配置通道的顺序为：先配置所有的直连网线通道，再配置应用所在网卡的通道，最后配置串口通道。

```
Enter number of heartbeat channels (minimum = 1) [1]: 4
You selected 4 channels
Information about channel 0:
Channel type: net or serial [net]:
Channel type set to: net
Enter hostname of cluster member test1 on heartbeat channel 0 [test1]:
hb11
Looking for host hb11 (may take a few seconds)...
Host hb11 found
Hostname corresponds to an interface on member 0
Channel name set to: hb11
Information about channel 1:
Channel type: net or serial [net]:
Channel type set to: net
Enter hostname of cluster member test1 on heartbeat channel 1: hb12
Looking for host hb12 (may take a few seconds)...
Host hb12 found
Hostname corresponds to an interface on member 0
Channel name set to: hb12
Information about channel 2:
Channel type: net or serial [net]:
Channel type set to: net
Enter hostname of cluster member test1 on heartbeat channel 2: test1
Looking for host test1 (may take a few seconds)...
Host test1 found
Hostname corresponds to an interface on member 0
Channel name set to: test1
Information about channel 3:
Channel type: net or serial [net]: serial
Channel type set to: serial
Enter device name: /dev/ttyS0
Device /dev/ttyS0 found and no getty running on it
Device name set to: /dev/ttyS0
```

4）配置本地节点的 power switch 和 watchdog 选项：

如果没有硬件电子开关设备，请输入 NONE。如果有硬件电子开关，输入相应的电子开关对应的类型，如：RSA、RPS10 或者是 APC。

软件级的 watchdog 可以用来保证 GreatTurbo Cluster Server 10 程序的健壮性但并不能保障操作系统的自动重启恢复； 如果有硬件 watchdog，请输入硬件

watchdog 对应的驱动模块的名字，如果没有硬件 watchdog，建议配置软件 watchdog，输入操作系统默认附带的软件 watchdog 的模块名字 softdog 即可。

```
Information about power switch connected to member 0
Specify one of the following switches (NONE/RSA/RPS10/APC) [NONE]: NONE
Power switch type set to NONE

Information about watchdog to member 0
Choose one of the following watchdog drivers: NONE/softdog/...) [NONE] :
softdog
```

5）配置对方节点信息

对方节点信息包括节点的机器名，heartbeat 的设置，watchdog driver 等。设置方法和本地完全一样。

```
----------------------------------------
Setting information for cluster member 1
----------------------------------------
Enter name of cluster member: test2
Looking for host test2 (may take a few seconds)...
Host test2 found
Cluster member name set to: test2

You previously selected 4 channels
Information about channel 0:
Channel type selected as net
Enter hostname of cluster member test2 on heartbeat channel 0: hb21
Looking for host hb21 (may take a few seconds)...
Host hb21 found
Channel name set to: hb21
Information about channel 1:
Channel type selected as net
Enter hostname of cluster member test2 on heartbeat channel 1: hb22
Looking for host hb22 (may take a few seconds)...
Host hb22 found
Channel name set to: hb22
Information about channel 2:
Channel type selected as net
Enter hostname of cluster member test2 on heartbeat channel 2: test2
Looking for host test2 (may take a few seconds)...
Host test2 found
```

```
Channel name set to: test2
Information about channel 3:
Channel type selected as serial
Enter device name [/dev/ttyS0]: /dev/ttyS0
Device name set to: /dev/ttyS0

Information about power switch connected to member 1
Specify one of the following switches (NONE/RSA/RPS10/APC) [NONE]: NONE
Power switch type set to NONE

Information about watchdog to member 1
Choose one of the following watchdog drivers(NONE/softdog/...) [NONE]:
softdog
```

6）配置磁盘心跳设备

利用共享 raw 磁盘分区作为心跳通道后，只要主备节点能够访问共享数据，就不会发生裂脑，从而有效的确保了共享数据的一致性。如果需要配置磁盘心跳，请选择 yes 继续配置。如果条件实在不具备，输入 no，然后回车。

```
----------------------------
Setting up raw disk heartbeat
----------------------------

Do you want add raw disk heartbeat device? (yes/no) [yes]:
The raw disk heartbeat device must be a raw disk device.
To enhance redundancy, need to configure two raw disk devices.

Enter the name of first raw disk heartbeat device [/dev/raw/raw1]:
/dev/raw/raw1

Enter the name of second raw disk heartbeat device [/dev/raw/raw2]:
/dev/raw/raw1
Now begin to initialize the raw disk heartbeat device, it will cause the
data in the devices lost.
Are you sure to initialize the raw devices which you input? (yes/no)
[yes]:yes
```

7）配置第三方 IP 地址

您需要配置两个节点都可以连接的第三方 IP 地址（要求能 ping 通，一般选择网关作为第三方 IP）。第三方 IP 能配置多个，以进一步提高系统高可用性。如果

要使配置的第三方 IP 生效，必须将两个节点的对外网络设置为备用心跳。**如果条件具备，建议您配置第三方 IP。**

如果条件实在不具备，输入 no，然后回车。

```
------------------------------
Setting up the third partner ip
------------------------------

Do you want cluster to determine network status? (yes/no) [yes]:

The IP address of third computer is needed to determine network status.
The third computer should be up all the time, so it is recommend to use
gateway IP address here. Please use IP address instead of domain name.
If you want input multiple third part IP, please use comma to separate.
For example, 192.168.0.1,192.168.0.3,192.168.0,10

Enter the IP address of third computer [172.16.76.1]:
     ……
```

8）配置负载均衡调度信息

首先需要输入两个调度节点的 ip 地址，然后选择一种调度方式。注意此时选择的调度方式必须和安装软件包时所选择的方式一致（请选择和 realserver 保持一致的调度方式）。如果选择 NAT 作为负载调度技术，那么还需要配置 NAT router 地址的相关信息。

```
----------------------------------------
Setting Load Balancer informations
----------------------------------------
Enter load balancer primary server IP [172.16.70.138]:
Looking for IP address 172.16.70.138 (may take a few seconds)...
IP address 172.16.70.138 found
IP address corresponds to an interface on member 0
Load balancer primary server IP set to: 172.16.70.138
Enter load balancer backup server IP [172.16.70.76]:
Looking for IP address 172.16.70.76 (may take a few seconds)...
IP address 172.16.70.76 found
Load balancer backup server IP set to: 172.16.70.76
Enter  load  balancer  network  forwarding  type(direct/nat/tunnel)
[direct]: nat
Load balancer network forwarding type set to: nat
Enter NAT router IP address: 172.16.70.76
```

```
NAT router IP address set to: 172.16.70.76
Enter NAT router netmask: 255.255.255.0
NAT router netmask set to: 255.255.255.0
Enter NAT router device(e.g. eth1:1): eth1:1
NAT router device set to: eth1:1
------------------------------------
The following choices will be saved:
------------------------------------
     ……
```

9）保存配置

如果以上内容全部配置完毕，member_config 将会询问是否保存改动。如果刚才的配置没有错误，请输入"yes"或直接回车；

然后程序还会询问是否运行"diskutil -I"来初始化配置文件，这里请选择"yes"。

注意：当 GreatTurbo Cluster Server 10 的守护进程正在运行时，请不要选择运行 diskutil -I 来初始化配置文件，以免造成不可预知的后果。

```
Save changes? yes/no [yes]:
Writing to output configuration file...done.
Changes have been saved to /etc/opt/cluster/cluster.conf.
----------------------------
Setting up Configure File
----------------------------
Run diskutil -I to set up the configure file now?
- Select 'yes' to clean up a previous install
- Select 'no' if you have just set them up on other member
  and have not started the cluster services on that member
- Select 'no' if you are running it on other cluster member
Warning: Do not run 'diskutil -I' on a running cluster, because it would
have severe consequences. yes/no [no]: yes

Saving configuration information to configure file: done

-----------------------------------------------------------------
Setup on this member is complete.  If errors have been reported,
correct them.

If you have not already set up the other cluster member, before
running member_config, invoke the following operation on the
```

```
other cluster member:
copy /etc/opt/cluster/cluster_raw.conf to another node with the same path.
cluster daemons on each cluster member by invoking the cluster start
script located in the System V init directory.  For example:
# /etc/rc.d/init.d/cluster start
```

2. 在对方节点上同步配置

配置完成之后，需要把配置文件/etc/opt/cluster/cluster_raw.conf 手工拷贝到另外一台机器的相同目录里(使用 scp 或者 ftp 均可)。

```
[root@test2 root]# scp test1:/etc/opt/cluster/cluster*
/etc/opt/cluster
```

3. 运行 GreatTurbo Cluster server 10

初始化完成之后，需要在两边节点分别运行 GreatTurbo Cluster Server 10。脚本/etc/init.d/cluster 可以用来启动 GreatTurbo Cluster Server 10。

```
[root@test1 root]# /etc/init.d/cluster start
-------- Performing cluster start ---------------------
Starting Turbo cluster...done.
---- Performing guiadmin server start ----
Starting guiadmin server...done.
---- Completed guiadmin server start ----
-------- Completed cluster start ---------------------
```

注意：请不要执行完/etc/init.d/cluster start 后立即执行/etc/init.d/cluster stop，请确认 GreatTurbo Cluster Server 10 启动完成后再执行/etc/init.d/cluster stop。

member_config 初始化完成之后，GreatTurbo Cluster Server 10 就可以正常运行了，只不过没有配置服务。用户可以参阅第十四章的检测方法来判断 GreatTurbo Cluster Server10 是否可以成功启动。

在配置服务之前，首先需要准备服务所需要的用户脚本和检测脚本。

第 4 章

用户脚本和检测脚本

在配置 GreatTurbo Cluster Server 10 的服务时，需要配置服务要使用的用户脚本和检测脚本。用户脚本也就是启动和停止脚本，用来启动和停止服务相关的应用程序。检测脚本在服务运行过程中用来检测服务的运行情况。用户脚本和检测脚本虽然并不是必需的，这要根据应用的实际情况来决定。如果需要用户脚本和检测脚本，那么在配置服务之前首先需要准备好用户脚本和检测脚本，并且测试用户脚本和检测脚本都正确无误后方可开始配置服务。

注意：对于 LB 服务，系统提供了默认脚本供用户使用，故本章讨论的脚本问题均针对 HA 服务，而不对 LB 服务使用的脚本进行详述。关于 LB 服务使用的脚本，具体参见第九、十、十一、十二章。

1. 用户脚本

用户脚本是一个 shell 的脚本，用来启动和停止 HA 服务相关的应用程序，一个服务只能配置一个用户脚本。

(1) 编写用户脚本

请参照下面的例子来编写用户脚本，在 GreatTurbo Cluster Server 10 安装光盘的 doc/examples 目录下也有用户脚本的例子，名字为 usrapp.sh。

在编写用户脚本时，需要注意以下的问题：

1) 脚本的参数必须包含对输入参数 start、stop 的相应处理，在启动部分加入用户所有程序的启动命令，在停止部分加入用户所有程序的停止命令。也就是说脚本的格式必须是规定的格式。请参照例子进行编写。

2) 请在脚本中加入 log 信息，以便发生故障时准确定位故障原因。

3) 脚本执行成功时要明确返回 0，失败时要明确返回 1。特别是当应用程序有标准的启动/停止脚本时，请确认该脚本在成功时是否返回 0，失败时是否返回 1，如果不是，请不要直接使用应用程序提供的脚本，请手动编写脚本。

4) 对于操作系统默认安装的一些应用的脚本，如 httpd，该脚本的返回值在启动/停止时会判断是否已经启动/停止了，在已经启动/停止时，该脚本返回值是 1，这不满足 GreatTurbo Cluster Server 10 的接口，需要将其改成

0。当然，用户自写的脚本在判断是否已经启动/停止时，也应该满足返回值为 0 的条件。

5) 如果启动和停止分别有多步操作，停止操作的执行顺序一定要和启动操作的执行顺序相反。

6) 用户脚本要有可执行权限。

(2) 测试用户脚本

用户脚本写好后，请务必通过测试检查用户脚本是否正确。

例如，用户脚本的名字为 usrapp.sh。

1) 启动

首先执行 usrapp.sh start。

然后 echo $? ,判断脚本的返回值是否为 0，0 表示脚本执行成功。

检查应用程序是否正常启动。

如果脚本的返回值为 0 并且应用程序正常启动了，表示用户脚本在启动时是正确的。

2) 停止

首先执行 usrapp.sh stop。

然后 echo $? ，判断脚本的返回值是否为 0，0 表示脚本执行成功。

检查应用程序是否正常停止。

如果脚本的返回值为 0 并且应用程序正常停止了，表示用户脚本在停止时是正确的。

(3) 用户脚本的例子

```
#!/bin/sh

# userscript sample

LOG_EMERG=0            # system is unusable
LOG_ALERT=1            # action must be taken immediately
LOG_CRIT=2         # critical conditions
LOG_ERR=3          # error conditions
LOG_WARNING=4      # warning conditions
LOG_NOTICE=5       # normal but significant condition
```

```
LOG_INFO=6        # informational
LOG_DEBUG=7           # debug-level messages

script_name=`basename $0`

culog()
{
        log_level=$1
        log_info=$2
        /opt/cluster/bin/culog -p $$ -n $script_name -s $log_level "$log_info"
}

case "$1" in
start)
    # start your application, put actual start actions here
    su - oracle -c "dbstart"
    if [ $? -eq 0 ]; then
        # should check application process here
        pmon=`ps -ef | egrep ora_pmon_$ORACLE_SID | grep -v grep`
        if [ "$pmon" = "" ];
        then
            culog $LOG_ERR " oracle database start failed, process not exist."
            # must return 1 here
            exit 1
        fi
        culog $LOG_INFO " dbstart succeeded."
    else
        culog $LOG_ERR " dbstart failed, ret=$?."
        # must return 1 here
        exit 1
    fi
    su - oracle -c "lsnrctl start"
    if [ $? -eq 0 ]; then
        culog $LOG_INFO " lsnrctl start succeeded."
        # must return 0 here
        exit 0
    else
        culog $LOG_ERR " lsnrctl start failed, ret=$?."
        # must return 1 here
        exit 1
    fi

     ;;

stop)
    #stop your application, put actual stop actions here
    su - oracle -c "lsnrctl stop"
    if [ $? -eq 0 ]; then
        culog $LOG_INFO " lsnrctl stop succeeded."
    else
        culog $LOG_ERR " lsnrctl stop failed, ret=$?."
```

```
        #must return 1 here
        exit 1
    fi
    su - oracle -c "dbshut"
    if [ $? -eq 0 ]; then
        pmon=`ps -ef | egrep ora_pmon_$ORACLE_SID | grep -v grep`
        if [ "$pmon" != "" ];
        then
            clulog $LOG_ERR " oracle database shut failed, process still exist."
            # must return 1 here
            exit 1
        fi
        clulog $LOG_INFO " dbshut succeeded."
        # must return 0 here
        exit 0
    else
        clulog $LOG_ERR " dbshut failed, ret=$?."
        # must return 1 here
        exit 1
    fi

    ;;

esac
```

2. 检测脚本

服务的检测脚本是一个 shell 的脚本，用来检测 HA 服务是否正常运行，一个 HA 服务只能配置一个检测脚本。

根据应用的实际情况，您可以自己编写检测脚本，也可以使用 GreatTurbo Cluster Server 10 自带的位于/opt/cluster/usercheck/目录下的 agent 脚本。

(1) 编写检测脚本

请参照下面的例子来编写检测脚本，在 GreatTurbo Cluster Server 10 安装光盘的 doc/examples 目录下也有检测脚本的例子，名字为 usrchk.sh。

在编写检测脚本时，需要注意以下的问题：

1) 请在脚本中加入 log 信息，以便发生故障时准确定位故障原因。为减小 log 文件的大小，检查成功时请不要打印 log 信息。

2) 脚本执行成功时返回 0，失败时返回非 0 值。当有多步检查时，请将每步检查失败时的返回值设置为不同的值。

3) 服务检测脚本的名称不要和待检测服务应用程序进程的名字相同。

4） 检测脚本要有可执行权限。

(2) 测试检测脚本

检测脚本写好后，请务必测试检测脚本是否正确。

例如，检测脚本的名字为 usrchk.sh。

1） 应用程序正常运行时

执行检测脚本 usrchk.sh。

然后 echo $? ，判断脚本的返回值是否为 0。

如果脚本的返回值为 0，表示检测脚本是正确的。

2） 应用程序运行不正常时

执行检测脚本 usrchk.sh。

然后 echo $? ，判断脚本的返回值是否为非 0 值。

如果脚本的返回值为非 0 值，表示检测脚本是正确的。

(3) 检测脚本的例子

```
#!/bin/sh

# checkscript sample

LOG_EMERG=0                 # system is unusable
LOG_ALERT=1                 # action must be taken immediately
LOG_CRIT=2              # critical conditions
LOG_ERR=3               # error conditions
LOG_WARNING=4           # warning conditions
LOG_NOTICE=5            # normal but significant condition
LOG_INFO=6              # informational
LOG_DEBUG=7                 # debug-level messages

script_name=`basename $0`

culog()
{
        log_level=$1
        log_info=$2
        /opt/cluster/bin/culog -p $$ -n $script_name -s $log_level "$log_info"
}
```

```
# check applications'process parts, you can add or delete part to check your actual
processes

/opt/cluster/usercheck/oracleCheck 172.16.74.127 1521
if [ $? -ne 0 ]
then
        # please add log information here
        clulog $LOG_ERR "oracleCheck failed, ret=$?."
        exit 1
fi

ps -ef | grep ora_pmon_ora10g | grep -v $0 | grep -v grep > /dev/null
if [ $? -ne 0 ]
then
        # please add log information here
        clulog $LOG_ERR "check oracle instance process failed."
        exit 2
fi

ps -ef | grep tnslsnr | grep -v $0 | grep -v grep > /dev/null
if [ $? -ne 0 ]
then
        # please add log information here
        clulog $LOG_ERR " check oracle listener process failed."
        exit 3
fi

# if all checks aren't wrong, then return 0
exit 0
```

脚本编写完成并检查无误后，就可以配置 HA 服务了。依照每个应用类型的不同，可以分别按照第五章、第六章、第七/八章的内容进行配置。

1）第一类 HA 应用的用户（需要使用共享磁盘阵列的应用）直接按照第五章进行配置。

2）第二类 HA 应用的用户（不需要使用共享磁盘阵列并且也不需要存取共享数据的应用）直接按照第六章进行配置。

3）第三类 HA 应用的用户（需要使用共享磁盘阵列但是不具备磁盘阵列硬件条件的应用）直接按照第七/八章进行配置。

第 5 章

配置 HA 服务--使用共享磁盘

GreatTurbo Cluster Server 10 在两个节点均正常启动后，就可以开始将用户的应用配置成高可用的服务了。本章说明如何在 GreatTurbo Cluster Server 10 里面配置使用共享磁盘阵列的服务，对服务进行配置的过程如下：

1) 配置服务的名字、优先节点
2) 配置服务的用户脚本
3) 配置服务的检测脚本
4) 配置服务的浮动 IP
5) 配置共享磁盘设备的信息
6) 设置是否 disable 服务

1. 配置工具

GreatTurbo Cluster Server 10 有两个配置工具：文本界面的 cluadmin 工具和图形界面的 guiadmin 工具。

这里主要介绍 cluadmin 的使用方法，guiadmin 的使用方法请参照第十三章。

/opt/cluster/bin/cluadmin 是文本界面的管理工具，使用它可以很方便地配置服务。cluadmin 类似 bash，可以使用 TAB 键进行命令补全。

注意：<u>如果您使用的是 Turbolinux 中文版，则在运行 cluadmin 之前，请先在 bash 下执行“unset LC_CTYPE”。</u>

2. 配置条件

请在 GreatTurbo Cluster Server 10 运行时执行 cluadmin 配置服务。

注意：<u>GreatTurbo Cluster Server 10 采用时间戳来判断两个节点的配置是否同步，所以不要轻易修改系统时间。如果确实需要修改系统时间，请先停止 GreatTurbo Cluster Server 10，修改完系统时间后，再重新启动 GreatTurbo Cluster Server 10。</u>

3. 配置命令

运行/opt/cluster/bin/cluadmin，然后在 cluadmin 中按下两次 TAB 键，会显示如下所示的所有命令。（[Tab] [Tab]表示连续按两下 Tab 键）

```
cluadmin>[Tab] [Tab]
apropos
clear
exit
help
cluster status
cluster monitor
cluster loglevel
cluster loglevel syncd
cluster loglevel svcmgr
cluster loglevel svccheck
cluster loglevel powerd
cluster loglevel heartbeat
cluster loglevel clumon
cluster heartbeat
cluster mail from
cluster mail to
cluster mail smtpserver
cluster mail level
cluster reload
cluster name
cluster edit
cluster backup
cluster restore
cluster saveas
cluster restorefrom
service add
service show state
service show config
service show services
service modify
service disable
service enable
service relocate
service delete
nbd add
nbd delete
nbd show
help apropos
help clear
help exit
help help
help cluster status
help cluster monitor
```

```
help cluster loglevel
help cluster reload
help cluster name
help cluster edit
help cluster backup
help cluster restore
help cluster saveas
help cluster restorefrom
help service add
help service show state
help service show config
help service show services
help service modify
help service disable
help service enable
help service relocate
help service delete
help nbd add
help nbd delete
help nbd show
```

4. 配置服务

利用 cluadmin 的 service add 命令添加服务。

(1)执行 cluadmin

```
[root@test1 root]# cluadmin
Sat Apr  9 17:48:57 CST 2005
…
```

(2)执行 service add（GreatTurbo Cluster Server 10 最多只支持 16 个 HA 服务）

```
cluadmin> service add
…
Currently defined services:
```

(3)输入服务的名字（如果有多个服务，服务的名字不能重复）

```
Service name: svc01
```

(4)输入服务的类别（只能选择 ha）

```
Service type (ha/lb): ha
```

(5)配置服务的优先节点

如果一个服务有优先节点，当优先节点启动 GreatTurbo Cluster Server 10 时或者优先节点的网卡故障恢复时，这个服务将会自动迁移到优先节点上运行。

优先节点缺省值为 None。直接回车表示选择 None，即不选择优先节点。否则输入优先节点的 hostname 名称。

```
Preferred member [None]: test1
```

(6)配置服务的用户脚本

服务的用户脚本是启动和停止应用程序的脚本。一个服务最多只能配置一个用户脚本。配置服务的用户脚本时，请输入全路径名。

```
User script (e.g., /usr/foo/script or None) [None]: /etc/init.d/
httpd
```

(7)配置服务检测脚本

“Check script”输入的脚本有两种选择方式，请根据应用选择其中的一种。

第一种方式是使用 GreatTurbo Cluster Server 10 自带的 agent 脚本，该脚本位于/opt/cluster/usercheck 目录下，使用 agent 脚本时，必须按规定格式输入，例如：/opt/cluster/usercheck/httpCheck 172.16.70.100 80。

第二种方式使用自己编写脚本。

配置服务的检测脚本时，必须输入全路径名。

配置服务检测脚本时还需要指定如下的参数：

➢ “Check interval”是检测服务的时间间隔，建议根据用户应用要求的切换时间设定合适的值，一般输入 5 秒，表示每隔 5 秒时间执行一次检测脚本。

➢ “Check timeout”是检测脚本执行的超时时间，执行用户检测脚本的超时时间，建议输入 30 秒。

➢ “Max error count”是允许服务连续错误的次数，一般输入 3 次，表示出错 3 次后，服务进行切换。

```
Do you want to add a check script to the service (yes/no/?) [no]:
```

```
yes
Check Script Information
Check script (e.g., "/opt/cluster/usercheck/httpCheck
172.16.69.200 80" or None) [None]:
/opt/cluster/usercheck/httpCheck 172.16.70.100 80
Check interval (in seconds) [None]: 5
Check timeout (in seconds) [None]: 30
Max error count [None]: 3
```

(8)配置服务的浮动 IP

服务可以绑定浮动 IP，浮动 IP 随着服务而浮动，也就是说这个 IP 所在的节点也就是服务所在的节点。一个服务最多可以配置 16 个浮动 IP。

注意：配置浮动 IP 时必须正确输入如下参数，注意随着 IP 地址的不同，相应的子网掩码以及广播地址的值需与之对应。

- “Net interface”是浮动 IP 绑定的网卡。
- “Netmask”是浮动 IP 的子网掩码。
- “Broadcast”是浮动 IP 的广播地址。

```
Do you want to add floating IP address to the service (yes/no/?)
[no]: yes

        IP Address Information

IP address: 172.16.70.100
Net interface [None]: eth0
Netmask (e.g. 255.255.255.0 or None) [None]: 255.255.255.0
Broadcast (e.g. X.Y.Z.255 or None) [None]: 172.16.71.255
Do you want to (a)dd, (m)odify, (d)elete or (s)how an IP address,
or are you (f)inished adding IP addresses: f
```

(9)配置共享磁盘设备的信息

因为我们在这里使用共享磁盘阵列，所以这里选择 yes。

GreatTurbo Cluster Server 10 最多可以支持 16 个磁盘设备。

我们支持两种类型的磁盘设备：一种是需要用户手动 mount 的磁盘设备；另外一种是不需要 mount 的裸设备。

对于第一种磁盘设备，我们必须按照下述要求输入相应参数。（以下例子假设用户需要 mount 的磁盘阵列的共享分区为/dev/sdb0）

必须指定的参数如下：

- “Device special file”是磁盘分区的设备名。
- “Check service device”指定是否检查磁盘设备，默认为 yes。
- “Device is SCSI”指定磁盘设备是否是 SCSI 设备，默认为 yes，请确认一下您的共享磁盘设备的类型，如果不是 SCSI 设备请选择 no。
- “Device check timeout”是磁盘检测的超时时间，默认为 120 秒。一般情况下选默认值就可以，如果磁盘负载比较大，请适当调大超时时间。
- “Filesystem type”是文件系统的类型。
- “Mount point”是 mount 的位置。
- “Mount options”是 mount 时的选项，一般用 rw, sync。
- “Forced unmount support”是指定 umount 时是否杀死该分区上运行的进程。
- “Device owner”是 mount 时的用户名，一般为 root。
- “Device group”是 mount 时的用户组名，一般为 root。
- “Device mode”是 mount 时的访问权限，一般为 755。

```
Do you want to add a disk device to the service (yes/no/?) [no]:
yes

        Disk Device Information

Device special file (e.g., /dev/sda1): /dev/sdb0
Check service device (yes/no/?) [yes]:
Device is SCSI (yes/no/?) [yes]:
Device check timeout (in seconds) [120]:
Filesystem type (e.g., ext2, reiserfs, ext3 or None): ext3
Mount point (e.g., /usr/mnt/service1) [None]: /mnt/userData
Mount options (e.g., rw,nosuid): rw
Forced unmount support (yes/no/?) [no]: yes
Device owner (e.g., root): root
Device group (e.g., root): root
Device mode (e.g., 755): 755
Do you want to (a)dd, (m)odify, (d)elete or (s)how devices, or
are you (f)inished adding devices: f
```

一次只能输入一个分区设备的信息，如果有多个分区设备，可以在结束提示时，不选择 f，而选择 a 继续增加分区设备。

对于第二种磁盘设备即裸设备，只需输入磁盘分区的设备名称，及指定磁盘检测的相关信息，其他信息必须全部按回车键略过；（以下例子假设裸设备对应的文件名是/dev/sdc）

```
Do you want to add a disk device to the service (yes/no/?) [no]:
yes

        Disk Device Information

Device special file (e.g., /dev/sda1): /dev/sdc
Check service device (yes/no/?) [yes]:
Device is SCSI (yes/no/?) [yes]:
Device check timeout (in seconds) [120]:
Filesystem type (e.g., ext2, reiserfs, ext3 or None):
Mount point (e.g., /usr/mnt/service1) [None]:
Mount options (e.g., rw,nosuid):
Forced unmount support (yes/no/?) [no]:
Device owner (e.g., root):
Device group (e.g., root):
Device mode (e.g., 755):
Do you want to (a)dd, (m)odify, (d)elete or (s)how devices, or
are you (f)inished adding devices: f
```

同样一次只能输入一个裸设备的名字，如果有多个裸设备，可以在结束提示时，不选择 f，而选择 a 继续增加分区设备。

(10) 设置启动服务的超时时间

建议此处直接回车，以选择服务默认启动超时时间 3600 秒。

```
Service start timeout (in seconds) [None]:
```

(11) 设置停止服务的超时时间

建议此处直接回车，以选择服务默认停止超时时间 3600 秒。

```
Service stop timeout (in seconds) [None]:
```

(12) 设置停止服务失败时是否 reboot 机器

如果设为“ yes ”，当服务停止失败时，为了释放服务的资源，将自动 reboot 机器。

如果设为“ no ”，当服务停止失败时，不会自动 reboot 机器，需要用户手动干预。

```
Reboot system if stop the service failed (yes/no/?) [yes]: yes
```

(13)　设置是否 disable 服务

如果设为“ yes ”，服务将不会立即被启动，只能以后由用户手动启动。
如果设为“ no ”，服务将立即被启动。一般选择 no，表示立即启动服务。

```
Disable service (yes/no/?) [no]: no
```

(14)　确认添加该服务

如果确认配置正确，请选择 yes 添加以上配置的服务。

```
Add svc01 service as shown? (yes/no/?) yes
Added svc01.
```

到这里，我们成功的添加了一个使用共享磁盘阵列的服务。如果用户还需要添加服务，那么仍然需要按照第四章末尾所述的内容来进行：选择第五、六、七、八章的其中一章来进行配置。

5. 修改服务

如果需要修改服务的配置，请执行 cluadmin 的 service modify 命令。

首先要求选择需要修改的服务，请输入列表中的服务的编号。

如果服务正在运行，将要求先停止服务，请输入 yes。

然后所有的操作都和添加服务相同。

```
cluadmin> service modify

  You will be prompted for information about the service.

  Enter a question mark (?) at a prompt to obtain help.

  Enter a colon (:) and a single-character command at a prompt to
do
  one of the following:

  c - Cancel and return to the top-level cluadmin command
  r - Restart to the initial prompt while keeping previous responses
  p - Proceed with the next prompt
```

```
  0) svc01
  c) cancel

Choose service to modify: 0
Modifying: svc01
Service is not disabled.  Disable it? (yes/no/?) yes
```

6. 删除服务

如果要删除服务，请执行 cluadmin 的 service delete 命令。

然后输入列表中的服务的编号即可。

```
cluadmin> service delete
  0) svc01
  c) cancel

Choose service to delete: 0
Deleting svc01, are you sure? (yes/no/?): yes
Service svc01 disabled
Svc01 deleted.
```

7. 启动服务

如果在添加服务时没有启动服务，需要手动启动服务，请执行 cluadmin 的 service enable 命令。

首先要求选择需要启动的服务，请输入列表中的服务的编号。

然后要求选择在哪个节点上启动服务，请输入节点的编号。

```
cluadmin> service enable
  0) svc01
  c) cancel

Choose service to enable: 0
Are you sure? (yes/no/?) yes
  0) test1
  1) test2
  c) cancel

Choose member: 0
```

```
Enabling svc01 on member test1. Service  enabled.
```

8. 停止服务

如果需要停止服务，请执行 cluadmin 的 service disable 命令。

首先选择要停止的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

```
cluadmin> service disable
  0) svc01
  c) cancel

Choose service to disable: 0
Are you sure? (yes/no/?) yes
Disabling svc01. Service  disabled.
```

9. 切换服务

如果需要把在一个节点上运行的服务切换到另一个节点上，请执行 cluadmin 的 service relocate 命令。

首先选择要切换的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

```
cluadmin> service relocate
  0) svc01
  c) cancel

Choose service to relocate: 0
Are you sure? (yes/no/?) yes
Relocating svc01. Service  relocated.
```

如果所有的服务添加完成，请直接到第十三章继续阅读随后内容。

第 6 章

配置 HA 服务--不使用共享磁盘

本章说明如何在 GreatTurbo Cluster Server 10 中配置不具有共享磁盘阵列的服务，配置服务的过程如下：

1) 配置服务的名字、优先节点
2) 配置服务的用户脚本
3) 配置服务的检测脚本
4) 配置服务的浮动 IP
5) 设置是否 disable 服务

1. 配置工具

GreatTurbo Cluster Server 10 有两个配置工具：文本界面的 cluadmin 工具和图形界面的 guiadmin 工具。

这里主要介绍 cluadmin 的使用方法，guiadmin 的使用方法请参照第十三章。

/opt/cluster/bin/cluadmin 是个文本界面的管理工具，使用它可以很方便的配置服务。cluadmin 类似 bash，可以使用 TAB 键进行命令补全。

注意：如果你使用的是 Turbolinux 中文版，则在运行 cluadmin 前，请先在 bash 下执行“unset LC_CTYPE”。

2. 配置条件

请在 GreatTurbo Cluster Server 10 运行时执行 cluadmin 配置服务。

注意：GreatTurbo Cluster Server 10 采用时间戳来判断两个节点的配置是否同步，所以不要轻易修改系统时间。如果确实需要修改系统时间，请先停止 GreatTurbo Cluster Server 10，修改完系统时间后，再重新启动 GreatTurbo Cluster Server 10。

3. 配置命令

运行/opt/cluster/bin/cluadmin，然后在 cluadmin 中按下两次 TAB 键，会显示如下所示的所有命令：

```
cluadmin>[Tab] [Tab]
```

```
apropos
clear
exit
help
cluster status
cluster monitor
cluster loglevel
cluster loglevel syncd
cluster loglevel svcmgr
cluster loglevel svccheck
cluster loglevel powerd
cluster loglevel heartbeat
cluster loglevel clumon
cluster heartbeat
cluster mail from
cluster mail to
cluster mail smtpserver
cluster mail level
cluster reload
cluster name
cluster edit
cluster backup
cluster restore
cluster saveas
cluster restorefrom
service add
service show state
service show config
service show services
service modify
service disable
service enable
service relocate
service delete
nbd add
nbd delete
nbd show
help apropos
help clear
help exit
help help
help cluster status
help cluster monitor
help cluster loglevel
```

```
help cluster reload
help cluster name
help cluster edit
help cluster backup
help cluster restore
help cluster saveas
help cluster restorefrom
help service add
help service show state
help service show config
help service show services
help service modify
help service disable
help service enable
help service relocate
help service delete
help nbd add
help nbd delete
help nbd show
```

4. 配置服务

利用 cluadmin 的 service add 命令添加服务。

(1) 执行 cluadmin

```
[root@test1 root]# cluadmin
Sat Apr  9 17:48:57 CST 2005
…
```

(2) 执行 service add（GreatTurbo Cluster Server 10 最多只支持 16 个 HA 服务）

```
cluadmin> service add
…
Currently defined services:
```

(3) 输入服务的名字（如果有多个服务，服务的名字不能重复）

```
Service name: svc01
```

(4) 输入服务的类别（只能选择 ha）

```
Service type (ha/lb): ha
```

(5) 配置服务的优先节点

如果一个服务有优先节点，当优先节点启动时，这个服务将会首先选择在优先节点上运行。

优先节点缺省值为 None。直接回车表示选择 None，即不选择优先节点。

```
Preferred member [None]: test1
```

(6) 配置服务的用户脚本

服务的用户脚本是启动和停止应用程序的脚本。一个服务最多只能配置一个用户脚本。配置服务的用户脚本时，必须输入全路径名。

```
User script (e.g., /usr/foo/script or None) [None]: /etc/init.d/
httpd
```

(7) 配置服务的检测脚本

“Check script”输入的脚本有两种选择方式，请根据应用选择其中的一种。

第一种方式是使用 GreatTurbo Cluster Server 10 自带的 agent 脚本，该脚本位于/opt/cluster/usercheck 目录下，使用 agent 脚本时，必须按规定格式输入，例如：/opt/cluster/usercheck/httpCheck 172.16.70.100 80。

第二种方式使用自己编写脚本。

配置服务的检测脚本时，必须输入全路径名。

配置服务检测脚本时还需要指定如下的参数：

- “Check interval”是检测服务的时间间隔，建议根据用户的应用要求的切换时间设定合适的值，一般输入 5 秒，表示每隔 5 秒时间执行一次检测脚本。
- “Check timeout”是检测脚本返回的超时时间，执行用户检测脚本的超时时间，建议输入 30 秒。
- “Max error count”是允许服务连续错误的次数，一般输入 3 次，表示出错 3 次后，服务进行切换。

```
Do you want to add a check script to the service (yes/no/?) [no]:
yes
Check Script Information
```

```
Check script (e.g., "/opt/cluster/usercheck/httpCheck
172.16.69.200 80" or None) [None]:
/opt/cluster/usercheck/httpCheck 172.16.70.100 80
Check interval (in seconds) [None]: 5
Check timeout (in seconds) [None]: 30
Max error count [None]: 3
```

(8)　配置服务的浮动 IP

服务可以绑定浮动 IP，浮动 IP 随着服务而浮动，也就是说这个 IP 所在的节点也就是服务所在的节点。一个服务最多可以配置 16 个浮动 IP。

注意：配置浮动 IP 时必须正确输入如下参数，注意随着 IP 地址的不同，相应的子网掩码以及广播地址的值需与之对应。

- “Net interface”是浮动 IP 绑定的网卡。
- “Netmask”是浮动 IP 的子网掩码。
- “Broadcast”是浮动 IP 的广播地址。

```
Do you want to add floating IP address to the service (yes/no/?)
[no]: yes

        IP Address Information

IP address: 172.16.70.100
Net interface [None]: eth0
Netmask (e.g. 255.255.255.0 or None) [None]: 255.255.255.0
Broadcast (e.g. X.Y.Z.255 or None) [None]: 172.16.71.255
Do you want to (a)dd, (m)odify, (d)elete or (s)how an IP address,
or are you (f)inished adding IP addresses: f
```

(9)　配置共享磁盘设备的信息

由于服务不使用共享磁盘设备，所有在此必须选择 no，直接回车。

```
Do you want to add a disk device to the service (yes/no/?) [no]:
```

(10)　设置启动服务的超时时间

建议此处直接回车，以选择服务默认启动超时时间 3600 秒。

```
Service start timeout (in seconds) [None]:
```

(11) 设置停止服务的超时时间

建议此处直接回车，以选择服务默认停止超时时间 3600 秒。

```
Service stop timeout (in seconds) [None]:
```

(12) 设置停止服务失败时是否 reboot 机器

如果设为“ yes ”，当服务停止失败时，为了释放服务的资源，将自动 reboot 机器。

如果设为“ no ”，当服务停止失败时，不会自动 reboot 机器，需要用户手动干预。

```
Reboot system if stop the service failed (yes/no/?) [yes]: yes
```

(13) 设置是否 disable 服务

如果设为“ yes ”，服务将不会立即被启动，只能以后由用户手动启动。

如果设为“ no ”，服务将立即被启动。一般选择 no，表示立即启动服务。

```
Disable service (yes/no/?) [no]: no
```

(14) 确认添加该服务

如果确认配置正确，请选择 yes 添加以上配置的服务。

```
Add svc01 service as shown? (yes/no/?) yes
Added svc01.
```

到这里，我们成功的添加了一个不使用共享磁盘阵列的服务。如果用户还需要添加服务，那么仍然需要按照第四章末尾所述的内容来进行：选择第五、六、七、八章的其中一章来进行配置。

5. 修改服务

如果需要修改服务的配置，请执行 cluadmin 的 service modify 命令。

首先要求选择需要修改的服务，请输入列表中的服务的编号。

如果服务正在运行，将要求先停止服务，请输入 yes。

然后所有的操作都和添加服务相同。

```
cluadmin> service modify

  You will be prompted for information about the service.

  Enter a question mark (?) at a prompt to obtain help.

  Enter a colon (:) and a single-character command at a prompt to
do
  one of the following:

  c - Cancel and return to the top-level cluadmin command
  r - Restart to the initial prompt while keeping previous responses
  p - Proceed with the next prompt



  0) svc01
  c) cancel

Choose service to modify: 0
Modifying: svc01
Service is not disabled.  Disable it? (yes/no/?) yes
```

6. 删除服务

如果要删除服务，请执行 cluadmin 的 service delete 命令。

然后输入列表中的服务的编号即可。

```
cluadmin> service delete
  0) svc01
  c) cancel

Choose service to delete: 0
Deleting svc01, are you sure? (yes/no/?): yes
Service svc01 disabled
Svc01 deleted.
```

7. 启动服务

如果在添加服务时没有启动服务，需要手动启动服务，请执行 cluadmin 的 service enable 命令。

首先要求选择需要启动的服务，请输入列表中的服务的编号。

然后要求选择在哪个节点上启动服务，请输入节点的编号。

```
cluadmin> service enable
  0) svc01
  c) cancel

Choose service to enable: 0
Are you sure? (yes/no/?) yes
  0) test1
  1) test2
  c) cancel

Choose member: 0
Enabling svc01 on member test1. Service  enabled.
```

8. 停止服务

如果需要停止服务，请执行 cluadmin 的 service disable 命令。

首先选择要停止的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

```
cluadmin> service disable
  0) svc01
  c) cancel

Choose service to disable: 0
Are you sure? (yes/no/?) yes
Disabling svc01. Service  disabled.
```

9. 切换服务

如果需要把在一个节点上运行的服务切换到另一个节点上，请执行 cluadmin 的 service relocate 命令。

首先选择要切换的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

```
cluadmin> service relocate
  0) svc01
  c) cancel

```

```
Choose service to relocate: 0
Are you sure? (yes/no/?) yes
Relocating svc01. Service  relocated.
```

如果所有的服务添加完成，请直接到第十三章继续阅读随后内容。

第 7 章

配置 HA 服务--使用磁盘镜像设备 kernel2.4

GreatTurbo Cluster Server 10 提供了虚拟的共享磁盘阵列的功能，使用磁盘镜像设备来保证应用数据的一致性。您可以使用节点本地普通的 IDE 或 SCSI 硬盘来构建磁盘镜像。

本章说明在 Linux kernel2.4 的环境下，如何在 GreatTurbo Cluster Server 10 中配置具有磁盘镜像设备的服务。

配置具有磁盘镜像设备的服务过程如下：

1) 安装磁盘镜像软件(drbd)。
2) 注册具有磁盘镜像功能的 license。
3) 配置磁盘镜像设备。
4) 添加使用磁盘镜像设备的服务。

1. 安装磁盘镜像软件(drbd)

在 GreatTurbo Cluster Server 10 安装时，会提示是否安装磁盘镜像软件，如果需要使用磁盘镜像软件，这时应当选择安装。

```
Test1:/mnt/cdrom # ./install_cluster
…
Do you want to use drbd? (y/n) [y]: y
…
```

选择了 y 后，安装程序会自动安装与当前系统相匹配的磁盘镜像软件(drbd)。

注意：

1) 磁盘镜像软件需要在两个节点上都要安装。
2) 请用安装程序自动安装磁盘镜像软件，不要手动安装磁盘镜像软件的 RPM 包。
3) 磁盘镜像软件中包含了 kernel 的模块，要求和 kernel 的版本相匹配，如果安装的过程中提示 drbd 的 kernel 模块和您当前系统的 kernel 版本不匹配，请更换成匹配的 kernel 或者获取更新的 drbd 安装包。

2．注册具有磁盘镜像功能的 license

使用 GreatTurbo Cluster Server 10 的磁盘镜像功能，需要注册 GreatTurbo Cluster Server 10 的带有磁盘镜像功能的 license。

License 的注册方法请参照【第 2 章 安装 GreatTurbo Cluster Server 10】。

在申请 GreatTurbo Cluster Server 10 的 license 时，请注意选择带 drbd 的 license。

3．配置磁盘镜像设备

(1)在两个节点上分别为磁盘镜像设备准备大小相同的磁盘分区。

(2)利用 cluadmin 工具配置磁盘镜像设备。

cluadmin 工具的详细使用方法请参照 GreatTurbo Cluster Server 10 的用户手册。

两个节点的 GreatTurbo Cluster Server 10 正常启动后，可以利用 cluadmin 的 nbd add 命令配置磁盘镜像设备。配置磁盘镜像设备的过程如下：

1) 执行 nbd add，输入 nbd ID。nbd ID 是镜像设备的编号，必须是 0-15 之间的整数。

```
Cluadmin> nbd add
Currently defined nbds:
Nbd ID(nbd identifier. e.g. 0): 0
```

2) 配置节点 A 的 nbd 信息。

```
Nbd configuration on member0:
name(nbd member ip address. e.g. 192.168.0.1): 192.168.1.11
device(block device used by nbd. e.g. /dev/hda14): /dev/hda7
port(port number used by nbd service. e.g. 8787): 8787
deviceNode(nbd block device. e.g. /dev/nb0): /dev/nb0
```

注意：

节点 A(member0)必须是执行 member_config 配置 GreatTurbo Cluster Server 10 时配置的第 1 个节点。

3) 配置节点 B 的 nbd 信息。

```
Nbd configuration on member1:
name(nbd member ip address. e.g. 192.168.0.1): 192.168.1.21
device(block device used by nbd. e.g. /dev/hda14): /dev/hda7
port(port number used by nbd service. e.g. 8787): 8787
```

```
deviceNode(
+nbd block device. e.g. /dev/nb0): /dev/nb0
```

注意：

节点 B(member1)必须是执行 member_config 配置 GreatTurbo Cluster Server 10 时配置的第 2 个节点。

4) 确认添加 nbd 设备。

```
Are you sure? (yes/no/?) y
Add nbd device 0
```

注意：

1) nbd 节点的 IP 必须采用双机直连的 IP。
2) 同一个 nbd 设备在两节点上的端口号要相同。
3) 磁盘镜像的设备名必须是/dev/nb#,并且同一个 nbd 设备在两节点上的设备名要相同。
4) 可以使用 nbd show 显示已添加的 nbd 设备；nbd delete 删除 nbd 设备。
5) 磁盘镜像设备配置完成后，将生成配置文件/etc/drbd.conf，请不要手动修改这个文件。

(3)启动磁盘镜像设备。

在两节点上分别执行命令 /etc/init.d/drbd start。

(4)初始化磁盘镜像设备。

1) 在节点 A 执行 control 0，使节点 A 成为第 0 号磁盘镜像设备的 Primary 节点。

```
[root@test1 root]# control 0
Setting 'drbd0' to Primary .. OK
datadisk: 'drbd0' activated
```

2) 查看磁盘镜像设备的状态。

```
[root@test1 root]# cat /proc/drbd
version: 0.6.8 (api:63/proto:62)

0: cs:Connected st:Primary/Secondary ns:0 nr:0 dw:0 dr:0
```

```
pe:0 ua:0
        NEEDS_SYNC
```

3) 如果磁盘镜像设备的状态显示为“NEEDS_SYNC”，需要同步两节点的磁盘镜像设备。

```
[root@test1 root]# datasync -l 0
Option l '0'
…
datadisk: 'drbd0' full synced!
```

4) 同步的过程中，可以查看进度。

```
[root@test1 root]# cat /proc/drbd
version: 0.6.8 (api:63/proto:62)

0: cs:SyncingAll st:Primary/Secondary ns:119100 nr:0 dw:0
dr:119108 pe:253 ua:0
        [==>.................]       sync'ed:       14.8%
(89611/104391)K
        finish: 0:06min speed: 14,780 (14,780) K/sec
```

5) 同步完成后，在 Primary 的节点上创建文件系统。例如，

```
[root@test1 root]# mkfs.ext3 /dev/nb0
```

(5)如果需要在磁盘镜像设备上安装应用程序的共享数据，需要执行本步；否则，跳过本步。

mount Primary 节点的磁盘镜像设备，安装应用程序的共享数据到磁盘镜像设备。例如，

```
[root@test1 root]# mount /dev/nb0 /mnt/hda7
…
```

4. 添加使用磁盘镜像设备的服务

利用 cluadmin 的 service add 命令添加服务。

(1) 执行 cluadmin

```
[root@test1 root]# cluadmin
Sat Apr  9 17:48:57 CST 2005
…
```

(2) 执行 service add

```
cluadmin> service add
…
Currently defined services:
```

(3) 输入服务的名字

```
Service name: svc01
```

(4) 输入服务的类别(只能选择 ha)

```
Service type (ha/lb): ha
```

(5) 配置服务的优先节点

如果一个服务有优先节点，当优先节点启动时，这个服务将会首先选择在优先节点上运行。

优先节点缺省值为 None。直接回车表示选择 None，即不选择优先节点。

```
Preferred member [None]: test1
```

(6) 配置服务的用户脚本

服务的用户脚本是启动和停止应用程序的脚本。一个服务最多只能配置一个用户脚本。配置服务的用户脚本时，请输入全路径名。

```
User script (e.g., /usr/foo/script or None) [None]:
/etc/init.d/httpd
```

(7) 配置服务检测脚本

“Check script”输入的脚本有两种选择方式，请根据应用选择其中的一种。

第一种方式是使用 GreatTurbo Cluster Server 10 自带的 agent 脚本，该脚本位于/opt/cluster/usercheck 目录下，使用 agent 脚本时，必须按规定格式输入，例如：/opt/cluster/usercheck/httpCheck 172.16.70.100 80。

第二种方式使用自己编写脚本。

配置服务的检测脚本时，必须输入全路径名。

配置服务检测脚本时需要指定如下的参数：

- “Check interval”是检测服务的时间间隔，缺省值是 3 秒。
- “Check timeout”是检测脚本返回的超时时间，缺省值是 30 秒。
- “Max error count”是允许服务连续错误的次数，缺省是 5 次。

```
Do you want to add a check script to the service (yes/no/?) [no]:
yes

        Check Script Information

Check script (e.g., "/opt/cluster/usercheck/httpCheck
172.16.69.200 80" or None) [None]:
/opt/cluster/usercheck/httpCheck 172.16.70.100 80
Check interval (in seconds) [None]: 5
Check timeout (in seconds) [None]: 10
Max error count [None]: 8
```

(8) 配置服务的浮动 IP

服务可以绑定浮动 IP，浮动 IP 随着服务而浮动，也就是说这个 IP 所在的节点也就是服务所在的节点。一个服务最多可以配置 16 个浮动 IP。

配置浮动 IP 时必须正确输入如下参数：

- “Net interface”是浮动 IP 绑定的网卡。
- “Netmask”是浮动 IP 的子网掩码。
- “Broadcast”是浮动 IP 的广播地址。

```
Do you want to add floating IP address to the service (yes/no/?)
[no]: yes

        IP Address Information

IP address: 172.16.70.100
Net interface [None]: eth0
Netmask (e.g. 255.255.255.0 or None) [None]: 255.255.255.0
Broadcast (e.g. X.Y.Z.255 or None) [None]: 172.16.70.255
Do you want to (a)dd, (m)odify, (d)elete or (s)how an IP address,
```

```
or are you (f)inished adding IP addresses: f
```

(9) 配置磁盘镜像设备

给服务配置磁盘镜像设备，一个服务最多可以配置 16 个磁盘镜像设备。

需要指定的参数如下：

- “Nbd device ID”指定磁盘镜像设备的 ID，所指定的磁盘镜像设备必须是前面配置过的。
- “Device special file”是磁盘分区的设备名，指定为磁盘镜像设备名。
- “Check service device”指定是否检测磁盘设备，这里请选择 no。

注意： **由于磁盘镜像设备自身有检测磁盘设备的功能，因此在配置磁盘镜像设备时请不要设置为检测磁盘设备。**

```
Do you want to add a nbd mirror disk device to the service
(yes/no/?) [no]: yes

        Nbd Information

Nbd device ID: 0
Do you want to (a)dd, (m)odify, (d)elete or (s)how nbds, or are
you (f)inished add nbds: f
Do you want to add a disk device to the service (yes/no/?) [no]:
yes

        Disk Device Information

Device special file (e.g., /dev/sda1): /dev/nb0
Check service device (yes/no/?) [yes]: no
```

(10) 配置磁盘镜像设备的 mount 信息

磁盘镜像设备的 mount 信息在配置服务的磁盘分区时设置，最多可以配置 16 个。

需要指定的参数如下：

- “Filesystem type”是文件系统的类型，请尽量选用日志文件系统。
- “Mount point”是 mount 的位置。
- “Mount options”是 mount 时的选项，一般用 rw, sync。

- “Forced unmount support”是指定 umount 时是否杀死该分区上运行的进程。
- “Device owner”是 mount 时的用户名，一般为 root。
- “Device group”是 mount 时的用户组名，一般为 root。
- “Device mode”是 mount 时的访问权限，一般为 755。

```
Filesystem type (e.g., ext2, reiserfs, ext3 or None): ext3
Mount point (e.g., /usr/mnt/service1) [None]: /mnt/hda7
Mount options (e.g., rw,nosuid): rw
Forced unmount support (yes/no/?) [no]: yes
Device owner (e.g., root): root
Device group (e.g., root): root
Device mode (e.g., 755): 755
Do you want to (a)dd, (m)odify, (d)elete or (s)how devices, or
are you (f)inished adding devices: f
```

(11) 设置启动服务的超时时间

启动服务的时间 = 启动 nbd 的时间 + mount 设备的时间 + 启动 IP 的时间 + 启动用户脚本的时间

启动服务的超时时间应当大于启动服务的时间，各个阶段的时间可以估算来设置，如果没有磁盘镜像设备，建议不设置启动服务的超时时间，采用缺省值。缺省值是 3600 秒(1 小时)。

由于启动 nbd 时，可能会进行两节点间数据的同步，因此比较花费时间。启动 nbd 的时间可以按如下的方法进行估算：

启动 nbd 的时间 = 磁盘分区大小/nbd 网络带宽 + 余量

当配置有多个 nbd 设备时，要分别计算。

例如：系统配置了 2 个 nbd 设备，分区大小分别为 100MB，500MB，使用同一个网络接口，网络带宽为 100Mbps。

启动 nbd 的时间 = (100*8/100 + 500*8/100)*2 + 100

= 196

考虑到 mount 设备的时间、启动 IP 的时间、启动用户脚本的时间，启动服务的超时时间可以设置为 300 秒。

```
Service start timeout (in seconds) [None]: 300
```

(12) 设置停止服务的超时时间

停止服务的时间 = 停止用户脚本的时间 + 停止 IP 的时间 + umount 设备的时间 + 停止 nbd 的时间

停止服务的超时时间应当大于停止服务的时间，各个阶段的时间可以估算来设置，如果没有磁盘镜像设备，建议不设置停止服务的超时时间，采用缺省值。缺省值是 3600 秒(1 小时)。

由于停止 nbd 时，可能会等待两节点间数据的同步，因此比较花费时间。

停止服务的超时时间的估算方法可以参考启动服务的超时时间的估算方法。

```
Service stop timeout (in seconds) [None]: 300
```

(13) 设置停止服务失败时是否 reboot 机器

如果设为“ yes ”，当服务停止失败时，为了释放服务的资源，将自动 reboot 机器。

如果设为“ no ”，当服务停止失败时，不会自动 reboot 机器，需要用户手动干预。

```
Reboot system if stop the service failed (yes/no/?) [yes]: yes
```

(14) 设置是否 disable 服务

如果设为“ yes ”，服务将不会立即被启动，只能由用户手动启动。

如果设为“ no ”，服务将立即被启动。

```
Disable service (yes/no/?) [no]: no
```

(15) 确认添加该服务

如果确认配置正确，请选择 yes 添加以上配置的服务。

```
Add svc01 service as shown? (yes/no/?) yes
Added svc01.
```

到这里，我们成功的添加了一个使用磁盘镜像设备的服务。如果用户还需要添加服务，那么仍然需要按照第四章末尾所述的内容来进行 选择第五、六、七、八章的其中一章来进行配置。

5. 修改服务

如果需要修改服务的配置，请执行 cluadmin 的 service modify 命令。

首先要求选择需要修改的服务，请输入列表中服务的编号。

如果服务正在运行，将要求先停止服务，请输入 yes。

然后所有的操作都和添加服务相同。

```
Cluadmin> service modify

  You will be prompted for information about the service.

  Enter a question mark (?) at a prompt to obtain help.

  Enter a colon (:) and a single-character command at a prompt to
do
  one of the following:

  c - Cancel and return to the top-level cluadmin command
  r - Restart to the initial prompt while keeping previous responses
  p - Proceed with the next prompt



  0) svc01
  c) cancel

Choose service to modify: 0
Modifying: svc01
Service is not disabled.  Disable it? (yes/no/?) yes
```

6. 删除服务

如果要删除服务，请执行 cluadmin 的 service delete 命令。

然后输入列表中的服务的编号即可。

```
cluadmin> service delete
  0) svc01
  c) cancel

Choose service to delete: 0
Deleting svc01, are you sure? (yes/no/?): yes
Service svc01 disabled
Svc01 deleted.
```

7. 启动服务

如果在添加服务时没有启动服务，需要手动启动服务，请执行 cluadmin 的 service enable 命令。

首先要求选择需要启动的服务，请输入列表中的服务的编号。

然后要求选择在哪个节点上启动服务，请输入节点的编号。

```
cluadmin> service enable
  0) svc01
  c) cancel

Choose service to enable: 0
Are you sure? (yes/no/?) yes
  0) test1
  1) test2
  c) cancel

Choose member: 0
Enabling svc01 on member test1. Service  enabled.
```

8. 停止服务

如果需要停止服务，请执行 cluadmin 的 service disable 命令。

首先选择要停止的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

```
cluadmin> service disable
  0) svc01
  c) cancel

Choose service to disable: 0
Are you sure? (yes/no/?) yes
Disabling svc01. Service  disabled.
```

9. 切换服务

如果需要把在一个节点上运行的服务切换到另一个节点上，请执行 cluadmin 的 service relocate 命令。

首先选择要切换的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

```
cluadmin> service relocate
  0) svc01
  c) cancel

Choose service to relocate: 0
Are you sure? (yes/no/?) yes
Relocating svc01. Service  relocated.
```

如果所有的服务添加完成，请到第十三章继续阅读随后内容。

第 8 章

配置 HA 服务--使用磁盘镜像设备 kernel2.6

GreatTurbo Cluster Server 10 提供了虚拟的共享磁盘阵列的功能，使用磁盘镜像设备来保证应用数据的一致性。您可以使用节点本地普通的 IDE 或 SCSI 硬盘来构建磁盘镜像。

本章说明在 Linux kernel2.6 的环境下，如何在 GreatTurbo Cluster Server 10 中配置具有磁盘镜像设备的服务。

配置具有磁盘镜像设备的服务过程如下：

1) 安装磁盘镜像软件(drbd)。
2) 注册具有磁盘镜像功能的 license。
3) 配置磁盘镜像设备。
4) 添加使用磁盘镜像设备的服务。

1. 安装磁盘镜像软件(drbd)

在 GreatTurbo Cluster Server 10 安装时，会提示是否安装磁盘镜像软件，如果需要使用磁盘镜像软件，这时应当选择安装。

```
Test1:/mnt/cdrom # ./install_cluster
…
Do you want to use drbd? (y/n) [y]: y
…
```

选择了 y 后，安装程序会自动安装与当前系统相匹配的磁盘镜像软件(drbd)。

注意：

1) 磁盘镜像软件需要在两个节点上都安装。
2) 请用安装程序自动安装磁盘镜像软件，不要手动安装磁盘镜像软件的 RPM 包。
3) 磁盘镜像软件中包含了 kernel 的模块，要求和 kernel 的版本相匹配，如果安装的过程中提示 drbd 的 kernel 模块和您当前系统的 kernel 版本不匹配，请更换成匹配的 kernel 或者获取更新的 drbd 安装包。

2. 注册具有磁盘镜像功能的 license

使用 GreatTurbo Cluster Server 10 的磁盘镜像功能，需要注册 GreatTurbo Cluster Server 10 的带有磁盘镜像功能的 license。

License 的注册方法请参照【第 2 章 安装 GreatTurbo Cluster Server 10】。

在申请 GreatTurbo Cluster Server 10 的 license 时，请注意选择带 drbd 的 license。

3. 配置磁盘镜像设备

(1)在两个节点上分别为磁盘镜像设备准备大小相同的磁盘分区。

(2)编辑磁盘镜像设备的配置文件。

在安装 GreatTurbo Cluster Server 10 时，会安装 drbd 的配置文件：/etc/drbd.conf，安装后的 drbd 的配置文件的内容如下所示：

```
#
# /etc/drbd.conf
#

# this is an example of drbd.conf
# please modify the following items according to your real environment.
# - hostname
# - device
# - disk
# - address
# if no particular reasons, you no need to modify other items.

resource drbd0 {
  protocol C;

  startup {
    wfc-timeout  30;
    degr-wfc-timeout 60;
  }

  syncer {
    rate 600M;
    group 0;
  }

  on hostname1 {
    device     /dev/drbd0;
    disk       /dev/hda6;
```

```
    address    192.168.0.1:7788;
    meta-disk  internal;
  }

  on hostname2 {
    device    /dev/drbd0;
    disk      /dev/hda6;
    address   192.168.0.2:7788;
    meta-disk internal;
  }
}
```

对/etc/drbd.conf 需要根据应用的实际环境编辑以下的内容：

- hostname，机器名。为 GreatTurbo Cluster Server 10 节点的机器名。
- device，设备名。为 drbd 设备名，可以选择/dev/drbd0，/dev/drbd1 等，两个节点所对应的 drbd 设备应当一致。
- disk，磁盘分区。为 drbd 设备所对应的物理磁盘分区。
- address，IP 地址和端口。为 drbd 设备通讯所用的网络及端口。Drbd 设备所用的网络最好是直连网络，不同的 drbd 设备所用的端口应当不同。

注意：

1) 关于 drbd 配置文件的详细说明请参照 GreatTurbo Cluster Server 10 用户手册。
2) 由于 meta-disk 需要占用 128MB 的磁盘空间，所以 drbd 所用的磁盘分区应当为 128MB 以上。
3) 由于 drbd 采用以太网络进行数据的传输和同步，所以 drbd 通讯所用的网络应当为直连网络，并且最好采用 bonding。
4) 如果要用多个 drbd 设备，请复制整个 resource 部分，并修改 resource 的名字，例如 reource1，然后再相应地对 hostname，device，disk 和 address 进行修改即可。

(3)启动磁盘镜像设备。

在两节点上分别执行命令 /etc/init.d/drbd start。

例如：

```
[root@test1 root]# /etc/init.d/drbd start
Starting DRBD resources:    [ d0 s0 n0 ].
```

(4)查看磁盘镜像设备的状态。

执行命令 cat /proc/drbd 来查看磁盘镜像设备的状态。

如果磁盘镜像设备的数据已经同步，状态显示如下：

```
[root@test1 root]# cat /proc/drbd
version: 0.7.11 (api:77/proto:74)
SVN Revision: 1912M build by root@qa3-127, 2005-08-25 10:20:28
 0: cs:Connected st:Secondary/Secondary ld:Consistent
    ns:0 nr:0 dw:0 dr:0 al:0 bm:92 lo:0 pe:0 ua:0 ap:0
```

如果磁盘镜像设备的数据没有同步，状态显示如下：

```
[root@test1 root]# cat /proc/drbd
version: 0.7.11 (api:77/proto:74)
SVN Revision: 1912M build by root@qa3-127, 2005-08-25 10:20:28
 0: cs:Connected st:Secondary/Secondary ld:Inconsistent
    ns:0 nr:0 dw:0 dr:0 al:0 bm:46 lo:0 pe:0 ua:0 ap:0
```

(5)使其中一个节点成为 Primary 节点。

如果状态显示为“ ld: Consistent ”，请在其中一个节点上执行“ drbdsetup /dev/drbd# primary ”。例如，

```
[root@test1 root]# drbdsetup /dev/drbd0 primary
[root@test1 root]# cat /proc/drbd
version: 0.7.11 (api:77/proto:74)
SVN Revision: 1912M build by root@qa3-127, 2005-08-25 10:20:28
 0: cs:Connected st:Primary/Secondary ld:Consistent
    ns:0 nr:366908 dw:366908 dr:0 al:0 bm:92 lo:0 pe:0 ua:0 ap:0
```

如果状态显示为“ ld: Inconsistent ”，请在其中一个节点上执行“ drbdsetup /dev/drbd# primary --do-what-I-say ”。这时，drbd 会自动进行同步。例如，

```
[root@test1 root]# drbdsetup /dev/drbd0 primary
--do-what-I-say
[root@test1 root]# cat /proc/drbd
version: 0.7.12 (api:77/proto:74)
SVN Revision: 1926M build by root@dev3-76, 2005-09-02 10:11:58
 0: cs:SyncSource st:Primary/Secondary ld:Consistent
    ns:152776 nr:0 dw:0 dr:160832 al:0 bm:1118 lo:756 pe:1418
ua:2014 ap:0
        [>...................] sync'ed:  2.5% (5776/5919)M
        finish: 0:02:00 speed: 49,036 (49,036) K/sec
```

(6)创建文件系统。

如果需要使用文件系统，则需要在 Primary 节点上创建文件系统。

```
[root@test1 root]# mkfs.ext3 /dev/drbd0
```

注意：

<u>文件系统只需要在 Primary 节点上创建，Secondary 节点会自动进行创建。</u>

(7)如果需要在磁盘镜像设备上安装应用程序的共享数据，需要执行本步操作；否则，跳过本步。

mount Primary 节点的磁盘镜像设备，安装应用程序的共享数据到磁盘镜像设备。例如，

```
[root@test1 root]# mount /dev/drbd0 /opt/oradata
…
```

(8)共享数据安装完成后，需要 umount 磁盘镜像设备。

```
[root@test1 root]# umount /dev/drbd0
```

(9)使 Primary 节点成为 Secondary 节点。

```
[root@test1 root]# drbdsetup /dev/drbd0 secondary
[root@test1 root]# cat /proc/drbd
version: 0.7.11 (api:77/proto:74)
SVN Revision: 1912M build by root@qa3-127, 2005-08-25 10:20:28
 0: cs:Connected st:Secondary/Secondary ld:Consistent
    ns:20348 nr:366908 dw:387256 dr:14468 al:47 bm:139 lo:0
pe:0 ua:0 ap:0
```

经过上述的操作以后，drbd 设备就准备好了，可以将 drbd 设备作为普通的磁盘进行使用，可以配置使用 drbd 设备的服务。

4. 添加使用磁盘镜像设备的服务

可以利用 cluadmin 工具或 guiadmin 工具来添加服务，这里主要介绍利用 cluadmin 工具配置的方法，guiadmin 工具的使用方法请参照第 9 章。

cluadmin 工具的详细使用方法请参照 GreatTurbo Cluster Server 10 的用户手册。

两个节点的 GreatTurbo Cluster Server 10 正常启动后，可以利用 cluadmin 的 service add 命令添加服务。

(1) 执行 cluadmin。

```
[root@test1 root]# cluadmin
Sat Apr  9 17:48:57 CST 2005
…
```

(2) 执行 service add。

```
cluadmin> service add
…
Currently defined services:
```

(3) 输入服务的名字。

```
Service name: svc01
```

(4) 输入服务的类别(只能选择 ha)。

```
Service type (ha/lb): ha
```

(5) 配置服务的优先节点。

如果一个服务有优先节点，当优先节点启动时，这个服务将会首先选择在优先节点上运行。

优先节点缺省值为 None。直接回车表示选择 None，即不选择优先节点。

```
Preferred member [None]: test1
```

(6) 配置服务的用户脚本。

服务的用户脚本是启动和停止应用程序的脚本。一个服务最多只能配置一个用户脚本。配置服务的用户脚本时，请输入全路径名。

```
User script (e.g., /usr/foo/script or None) [None]:
/etc/init.d/httpd
```

(7) 配置服务检测脚本。

“Check script”输入的脚本有两种选择方式，请根据应用选择其中的一种。

第一种方式是使用 GreatTurbo Cluster Server 10 自带的 agent 脚本，该脚本位于/opt/cluster/usercheck 目录下，使用 agent 脚本时，必须按规定格式输入，例如：/opt/cluster/usercheck/httpCheck 172.16.70.100 80。

第二种方式使用自己编写脚本。

配置服务的检测脚本时，必须输入全路径名。

配置服务检测脚本时需要指定如下的参数：

- “Check interval”是检测服务的时间间隔，缺省值是 3 秒。
- “Check timeout”是检测脚本返回的超时时间，缺省值是 30 秒。
- “Max error count”是允许服务连续错误的次数，缺省是 5 次。

```
Do you want to add a check script to the service (yes/no/?) [no]:
yes

        Check Script Information

Check script (e.g., "/opt/cluster/usercheck/httpCheck
172.16.69.200 80" or None) [None]:
/opt/cluster/usercheck/httpCheck 172.16.70.100 80
Check interval (in seconds) [None]: 5
Check timeout (in seconds) [None]: 10
Max error count [None]: 8
```

(8) 配置服务的浮动 IP。

服务可以绑定浮动 IP，浮动 IP 随着服务而浮动，也就是说这个 IP 所在的节点也就是服务所在的节点。一个服务最多可以配置 16 个浮动 IP。

配置浮动 IP 时必须正确输入如下参数：

- “Net interface”是浮动 IP 绑定的网卡。
- “Netmask”是浮动 IP 的子网掩码。
- “Broadcast”是浮动 IP 的广播地址。

```
Do you want to add floating IP address to the service (yes/no/?)
[no]: yes

        IP Address Information

IP address: 172.16.70.100
Net interface [None]: eth0
Netmask (e.g. 255.255.255.0 or None) [None]: 255.255.255.0
```

```
Broadcast (e.g. X.Y.Z.255 or None) [None]: 172.16.70.255
Do you want to (a)dd, (m)odify, (d)elete or (s)how an IP address,
or are you (f)inished adding IP addresses: f
```

(9) 配置磁盘设备。

在配置磁盘设备时，请指定相应的 drbd 设备。例如：/dev/drbd0。一个服务最多可以配置 16 个磁盘设备。

```
Do you want to add a disk device to the service (yes/no/?) [no]:
yes

        Disk Device Information

Device special file (e.g., /dev/sda1): /dev/drbd0
```

(10)配置磁盘设备的 mount 信息。

需要指定的参数如下：

- “Filesystem type”是文件系统的类型，请尽量选用日志文件系统。
- “Mount point”是 mount 的位置。
- “Mount options”是 mount 时的选项，一般用 rw, sync。
- “Forced unmount support”是指定 umount 时是否杀死该分区上运行的进程。
- “Device owner”是 mount 时的用户名，一般为 root。
- “Device group”是 mount 时的用户组名，一般为 root。
- “Device mode”是 mount 时的访问权限，一般为 755。

```
Filesystem type (e.g., ext2, reiserfs, ext3 or None): ext3
Mount point (e.g., /usr/mnt/service1) [None]: /mnt/hda7
Mount options (e.g., rw,nosuid): rw
Forced unmount support (yes/no/?) [no]: yes
Device owner (e.g., root): root
Device group (e.g., root): root
Device mode (e.g., 755): 755
Do you want to (a)dd, (m)odify, (d)elete or (s)how devices, or
are you (f)inished adding devices: f
```

(11)设置启动服务的超时时间。

启动服务的时间 = 启动 drbd 的时间 + mount 设备的时间 + 启动 IP 的时间 + 启动用户脚本的时间

启动服务的超时时间应当大于启动服务的时间，各个阶段的时间可以估算来设置，如果没有磁盘镜像设备，建议不设置启动服务的超时时间，采用缺省值。缺省值是 3600 秒(1 小时)。

由于启动 drbd 时，可能会进行两节点间数据的同步，因此比较花费时间。启动 drbd 的时间可以按如下的方法进行估算：

启动 drbd 的时间 = 磁盘分区大小/drbd 网络带宽 + 余量

当配置有多个 drbd 设备时，要分别计算。

例如：系统配置了 2 个 drbd 设备，分区大小分别为 500MB，1GB，使用同一个网络接口，网络带宽为 100Mbps。

启动 nbd 的时间 = (500*8/100 + 1024*8/100)*2 + 100

= 344

考虑到 mount 设备的时间、启动 IP 的时间、启动用户脚本的时间，启动服务的超时时间可以设置为 500 秒。

```
Service start timeout (in seconds) [None]: 500
```

(12)设置停止服务的超时时间。

停止服务的时间 = 停止用户脚本的时间 + 停止 IP 的时间 + umount 设备的时间 + 停止 drbd 的时间

停止服务的超时时间应当大于停止服务的时间，各个阶段的时间可以估算来设置，如果没有磁盘镜像设备，建议不设置停止服务的超时时间，采用缺省值。缺省值是 3600 秒(1 小时)。

由于停止 drbd 时，可能会等待两节点间数据的同步，因此比较花费时间。

停止服务的超时时间的估算方法可以参考启动服务的超时时间的估算方法。

```
Service stop timeout (in seconds) [None]: 500
```

(13)设置停止服务失败时是否 reboot 机器。

如果设为“yes”，当服务停止失败时，为了释放服务的资源，将自动 reboot 机器。

如果设为“no”，当服务停止失败时，不会自动 reboot 机器，需要用户手动干预。

```
Reboot system if stop the service failed (yes/no/?) [yes]: yes
```

(14)设置是否 disable 服务。

如果设为“ yes ”，服务将不会立即被启动，只能由用户手动启动。
如果设为“ no ”，服务将立即被启动。

```
Disable service (yes/no/?) [no]: no
```

(15)确认添加该服务。

如果确认配置正确，请选择 yes 添加以上配置的服务。

```
Add svc01 service as shown? (yes/no/?) yes
Added svc01.
```

到这里，我们成功的添加了一个使用磁盘镜像设备的服务。如果用户还需要添加服务，那么仍然需要按照第四章末尾所述的内容来进行 :选择第五、六、七、八章的其中一章来进行配置。

5. 修改服务

如果需要修改服务的配置，请执行 cluadmin 的 service modify 命令。

首先要求选择需要修改的服务，请输入列表中服务的编号。

如果服务正在运行，将要求先停止服务，请输入 yes。

然后所有的操作都和添加服务相同。

```
Cluadmin> service modify

  You will be prompted for information about the service.

  Enter a question mark (?) at a prompt to obtain help.

  Enter a colon (:) and a single-character command at a prompt to
do
  one of the following:

  c - Cancel and return to the top-level cluadmin command
  r - Restart to the initial prompt while keeping previous responses
  p - Proceed with the next prompt
```

```
  0) svc01
  c) cancel

Choose service to modify: 0
Modifying: svc01
Service is not disabled.  Disable it? (yes/no/?) yes
```

6. 删除服务

如果要删除服务，请执行 cluadmin 的 service delete 命令。

然后输入列表中的服务的编号即可。

```
cluadmin> service delete
  0) svc01
  c) cancel

Choose service to delete: 0
Deleting svc01, are you sure? (yes/no/?): yes
Service svc01 disabled
Svc01 deleted.
```

7. 启动服务

如果在添加服务时没有启动服务，需要手动启动服务，请执行 cluadmin 的 service enable 命令。

首先要求选择需要启动的服务，请输入列表中的服务的编号。

然后要求选择在哪个节点上启动服务，请输入节点的编号。

```
cluadmin> service enable
  0) svc01
  c) cancel

Choose service to enable: 0
Are you sure? (yes/no/?) yes
  0) test1
  1) test2
  c) cancel

Choose member: 0
Enabling svc01 on member test1. Service  enabled.
```

8. 停止服务

如果需要停止服务，请执行 cluadmin 的 service disable 命令。

首先选择要停止的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

```
cluadmin> service disable
  0) svc01
  c) cancel

Choose service to disable: 0
Are you sure? (yes/no/?) yes
Disabling svc01. Service  disabled.
```

9. 切换服务

如果需要把在一个节点上运行的服务切换到另一个节点上，请执行 cluadmin 的 service relocate 命令。

首先选择要切换的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

```
cluadmin> service relocate
  0) svc01
  c) cancel

Choose service to relocate: 0
Are you sure? (yes/no/?) yes
Relocating svc01. Service  relocated.
```

如果所有的服务添加完成，请到第十三章继续阅读随后内容。

第 9 章

配置 LB 服务--DR 方式

配置服务前要求用户在安装软件和初始化过程中均选择 dr 作为负载均衡技术。

GreatTurbo Cluster Server 10 在调度节点正常启动后，就可以开始将用户的应用配置成负载均衡的服务了。本章说明如何配置使用 dr 方式的服务，对服务进行配置的过程如下：

7) 配置 realserver 环境
8) 配置 director 环境
9) 在 director 上配置服务

1. 配置工具

GreatTurbo Cluster Server 10 使用 cluadmin 进行配置服务。

/opt/cluster/bin/cluadmin 是文本界面的管理工具，使用它可以很方便地配置服务。cluadmin 类似 bash，可以使用 TAB 键进行命令补全。

注意：如果您使用的是 Turbolinux 中文版，则在运行 cluadmin 之前，请先在 bash 下执行“unset LC_CTYPE”。

2. 配置条件

- 请在 GreatTurbo Cluster Server 10 运行时执行 cluadmin 配置服务。

  注意：GreatTurbo Cluster Server 10 采用时间戳来判断两个调度节点的配置是否同步，所以不要轻易修改系统时间。如果确实需要修改系统时间，请先停止 GreatTurbo Cluster Server 10，修改完系统时间后，再重新启动 GreatTurbo Cluster Server 10。
- 因为 dr 工作方式需要使用防火墙（iptables）作为转发工具，所以需要保证在作 realserver 的机器上安装 iptables 软件。
- 调度节点和所有 realserver 都必须在物理上有一个网卡通过不分断的局域网相连，如通过高速的交换机或者 HUB 相连。

3. 配置 realserver 环境

1) 以下 2-4 步操作要求在所有的 realserver 上都执行。
2) 根据业务需要修改/etc/init.d/lbrealserver-dr 文件。

  - ✧ 如果 realserver 上需要运行的服务没有在机器启动后自动运行,那么需要在上述文件 start、stop 部分的指定位置添加该服务的启动和停止脚本。
  - ✧ 如果 realserver 上运行的服务在机器启动后自动运行,那么不需要修改上述文件。

3) 根据 realserver 上运行的服务个数及配置，修改 /opt/cluster/lb/lbrealserver-conf-dr 文件。

   i. 如果 realserver 上运行 n 个服务，那么需要在 start、stop 部分复制 n-1 个 iptables 项。然后根据服务具体情况修改每一项的 ip 地址、dport、--to-port 值。如果仅有一个服务，那么只需要修改脚本中已经存在的 iptalbes 项。

   ii. 例，文件中 start 部分有如下项：

   iptables -t nat -A PREROUTING -p tcp -d 172.16.70.100 --dport 22 -j REDIRECT --to-port 22

   - ✧ 修改其中的 ip 地址“172.16.70.100”为服务的实际浮动 ip 地址。
   - ✧ 把 dport 后和--to-port 后的端口号“22”修改为服务实际占用的端口。
   - ✧ 根据上述内容，在 stop 部分作相应的修改，清除所设置的防火墙规则。

4) 执行“/etc/init.d/lbrealserver-dr start”命令，将该节点配置成一个 realserver。如果用户从集群中删掉该节点，那么需要执行“/etc/init.c/lbrealserver-dr stop”命令恢复原来设置。

5) 如果用户在已经运行的 GreatTurbo Cluster Server 10 系统上添加 LB 服务，那么需要在每个 realserver 上先后进行如下操作。
   - ✧ 执行/etc/init.d/lbrealserver-dr stop；
   - ✧ 修改/opt/cluster/lb/lbrealserver-conf-dr 文件；
   - ✧ 执行/etc/init.d/lbrealserver-dr start

4. 配置调度节点环境

1) 如果调度节点（director）没有作为 realserver 使用，那么 director 两节点不需要其他配置。

2) 如果 director 同时作为 realserver 使用，那么需要按照普通 realserver 配置一样首先配置环境。具体配置方法参见“配置 realsever 环境”

5. 配置命令

运行/opt/cluster/bin/cluadmin，然后在 cluadmin 中按下两次 TAB 键，会显示如下所示的所有命令。（[Tab] [Tab]表示连续按两下 Tab 键）

```
cluadmin>[Tab] [Tab]
apropos
clear
```

```
exit
help
cluster status
cluster monitor
cluster loglevel
cluster loglevel syncd
cluster loglevel svcmgr
cluster loglevel svccheck
cluster loglevel powerd
cluster loglevel heartbeat
cluster loglevel clumon
cluster heartbeat
cluster mail from
cluster mail to
cluster mail smtpserver
cluster mail level
cluster reload
cluster name
cluster edit
cluster backup
cluster restore
cluster saveas
cluster restorefrom
service add
service show state
service show config
service show services
service modify
service disable
service enable
service relocate
service delete
nbd add
nbd delete
nbd show
help apropos
help clear
help exit
help help
help cluster status
help cluster monitor
help cluster loglevel
help cluster reload
help cluster name
```

```
help cluster edit
help cluster backup
help cluster restore
help cluster saveas
help cluster restorefrom
help service add
help service show state
help service show config
help service show services
help service modify
help service disable
help service enable
help service relocate
help service delete
help nbd add
help nbd delete
help nbd show
```

6. 配置单端口服务

1) 执行 cluadmin

```
[root@test1 root]# cluadmin
Sat May  9 17:48:57 CST 2006
…
```

2) 执行 service add（GreatTurbo Cluster Server 10 最多只支持 256 个服务）

```
cluadmin> service add
…
Currently defined services:
```

3) 输入服务的名字（如果有多个服务，服务的名字不能重复）

```
Service name: svc01
```

4) 输入服务的类别（只能选择 lb）

```
Service type (ha/lb): lb
```

5） 配置服务的优先节点

如果服务有优先节点，当优先节点启动 LB 时或者优先节点的网卡故障恢复时，这个服务将会自动迁移到优先节点上运行。

优先节点缺省值为 None。直接回车表示选择 None，即不选择优先节点。否则输入优先节点的 hostname 名称。

如果配置优先节点，那么所有服务必须选择同一优先节点。

```
Preferred member [None]: test1
```

6） 配置服务的用户脚本

服务的用户脚本是启动和停止应用程序的脚本。配置服务的用户脚本时，请输入全路径名。所有 lb 服务的用户脚本需要保持相同。

如果两个调度节点同时作为 realserver 使用，那么使用 /opt/cluster/bin/lbdirector-real-dr 作为用户脚本。

```
User script (required, e.g., /opt/cluster/bin/lbdirector-dr):
/opt/cluster/bin/lbdirector-dr
```

7） 配置服务检测脚本

所有 lb 服务只能使用同一指定脚本：/opt/cluster/bin/lbcheck

配置服务检测脚本时还需要指定如下的参数：

- “Check interval”是检测服务的时间间隔，建议根据用户应用要求的切换时间设定合适的值，一般输入 5 秒，表示每隔 5 秒时间执行一次检测脚本。
- “Check timeout”是检测脚本执行的超时时间，执行用户检测脚本的超时时间，建议输入 30 秒。
- “Max error count”是允许服务连续错误的次数，一般输入 3 次，表示出错 3 次后，服务进行切换。

```
Do you want to add a check script to the service (yes/no/?) [no]:
yes
Check Script Information
Check script (e.g., "/opt/cluster/bin/lbcheck" or None) [None]:
/opt/cluster/bin/lbcheck
Check interval (in seconds) [None]: 5
Check timeout (in seconds) [None]: 30
Max error count [None]: 3
```

8） 配置服务的端口号

对于单端口服务只需要输入实际服务运行的端口号即可。如 http 为 80，ssh 为 22。

```
Application port(e.g., 80 or None) [None]: 22
```

9) 配置服务的协议

```
Protocol(tcp/udp) [tcp]: tcp
```

10) 配置服务的浮动 IP

服务可以绑定浮动 IP，浮动 IP 随着服务而浮动，也就是说这个 IP 所在的节点也就是服务所在的节点。

注意：配置浮动 IP 时必须正确输入如下参数，注意随着 IP 地址的不同，相应的子网掩码需与之对应。

- “address”是浮动 IP 的地址。
- “netmask”是浮动 IP 的子网掩码。
- “device” 是浮动 IP 所要绑定的网卡。

```
Virtual IP address(required): 172.16.71.111
Virtual IP network mask(optional) [None]: 255.255.255.0
Virtual IP device(required, e.g., eth0:1): eth0:1
```

11) 配置 firewall mark

单端口服务不需要配置，直接回车即可。

```
Firewall mark [None]:
```

12) 配置真实服务器重新加入集群的时间

建议使用默认值

```
Re-entry time(in seconds) [15]:
```

13) 配置检测 realserver 的时间间隔

建议使用默认值

```
Service timeout(in seconds) [6]:
```

14) 配置服务连接表选项

建议配置为 yes，表示当真实服务器重新加入集群内要把已经存在的连接表清空，避免新加入的服务器过度负载。

```
Quiesce server(yes/no) [no]:yes
```

15) 配置 realserver 的负载监视工具

调度节点能够使用 rup 或 ruptime 来监视各个真正服务器的载量。如果选择了 rup，每个真正服务器就必须运行 rstatd 服务。如果选择了 ruptime，每个真正服务器就必须运行 rwhod 服务。

注意： 载量监视和负载均衡不同。当与加权的调度算法综合使用时，能够导致难以预测的调度行为。还有，如果你要使用载量监视，群集中的真正服务器就必须是 Linux 机器。

```
Load monitoring tool(rup/ruptime/None) [None]:
```

16) 配置调度算法

根据服务的具体要求可以选择一种调度算法。

```
Scheduling(rr/wrr/lc/wlc/lblc/lblcr/dh/sh) [wlc]:
```

17) 配置持续连接时间

单端口服务不需要配置此项，回车即可。

```
Persistence(in seconds) [None]:
```

18) 配置 realserver 上服务的监视脚本

这里输入监视 realserver 上服务的脚本。我们提供了一些监视脚本，可以用来监视大多数常见服务；如果没有提供指定服务的脚本，您可以参照其他监视脚本自己编写或者选择 None。如果选择使用脚本检测 realserver，注意要在脚本路径后面使用特殊的“%h”符号。调度节点的 Nanny 进程会在调度该脚本时，使用 realserver 的真实 ip 地址代替%h。

```
Sending program(e.g., "/opt/cluster/bin/lbsend-http %h" or None) [None]: /opt/cluster/bin/lbsend-ssh %h
```

19) 配置 realserver 的回复文本

如果使用我们提供的脚本，那么要输入“OK”；如果使用其它方式，那么 expect 项根据具体脚本情况填写。

```
Expect string(e.g., "OK" or None) [None]: OK
```

20) 配置回复文本属性

如果回复文本时正则表达式，则此处选择 yes；否则选择 no

```
Treat expect string as a regular expression(yes/no) [no]: no
```

21) 配置 realserver

一个服务可以配置多个 realserver，调度器根据调度算法，将客户的请求调度到不同的 realserver 上。每个 realserver 都需要配置如下几项：

- “Server name”是一个用于真正服务器的描述性名称。
- “Server address”是真正服务器的 IP 地址。
- “Server weight” 是一个表明和集群内其它主机相比而言的主机能力的整数值。
- “Server active” 表示真实服务器是否启用。

```
Do you want to add server to the service (yes/no/?) [no]: y

        Server Information

Server name(required): server1
Server address(required): 172.16.70.211
Server weight [None]: 1
Server active(yes/no) [yes]: yes
Do you want to (a)dd, (m)odify, (d)elete or (s)how a server, or
are you (f)inished adding servers: f
```

22) 设置停止服务失败时是否 reboot 机器

如果设为“yes”，当服务停止失败时，为了释放服务的资源，将自动 reboot 机器。

如果设为“no”，当服务停止失败时，不会自动 reboot 机器，需要用户手动干预。

```
Reboot system if stop the service failed (yes/no/?) [yes]: yes
```

23) 设置是否 disable 服务

如果设为“yes”，服务将不会立即被启动，只能以后由用户手动启动。
如果设为“no”，服务将立即被启动。一般选择 no，表示立即启动服务。

```
Disable service (yes/no/?) [no]: no
```

24) 保存服务配置

选择 yes

```
Add test service as shown? (yes/no/?) yes
```

到这里，我们成功的添加了一个单端口的服务。如果用户还需要添加服务，那么重复上述过程。

7. 修改服务

如果需要修改服务的配置，请在 cluadmin 下执行 service modify 命令。

首先要求选择需要修改的服务，请输入列表中的服务的编号。

如果服务正在运行，将要求先停止服务，请输入 yes。

然后所有的操作都和添加服务相同。

```
cluadmin> service modify

  You will be prompted for information about the service.

  Enter a question mark (?) at a prompt to obtain help.

  Enter a colon (:) and a single-character command at a prompt to
do
  one of the following:

  c - Cancel and return to the top-level cluadmin command
  r - Restart to the initial prompt while keeping previous responses
  p - Proceed with the next prompt
```

```
  0) svc01
  c) cancel

Choose service to modify: 0
Modifying: svc01
Service is not disabled.  Disable it? (yes/no/?) yes
```

8. 删除服务

如果要删除服务，请执行 cluadmin 的 service delete 命令。

然后输入列表中的服务的编号即可。

```
cluadmin> service delete
  0) svc01
  c) cancel

Choose service to delete: 0
Deleting svc01, are you sure? (yes/no/?): yes
Service svc01 disabled
Svc01 deleted.
```

9. 启动服务

如果在添加服务时没有启动服务，需要手动启动服务，请执行 cluadmin 的 service enable 命令。

首先要求选择需要启动的服务，请输入列表中的服务的编号。

然后要求选择在哪个节点上启动服务，请输入节点的编号。

注意：启动一个服务的同时，所有的服务都将被启动。

```
cluadmin> service enable
  0) svc01
  c) cancel

Choose service to enable: 0
Are you sure? (yes/no/?) yes
  0) test1
  1) test2
  c) cancel
```

```
Choose member: 0
Enabling svc01 on member test1. Service  enabled.
```

10. 停止服务

如果需要停止服务，请执行 cluadmin 的 service disable 命令。

首先选择要停止的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

注意： 停止一个服务的同时，所有的服务都将被停止。

```
cluadmin> service disable
  0) svc01
  c) cancel

Choose service to disable: 0
Are you sure? (yes/no/?) yes
Disabling svc01. Service  disabled.
```

11. 切换服务

如果需要把在一个节点上运行的服务切换到另一个节点上，请执行 cluadmin 的 service relocate 命令。

首先选择要切换的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

注意：切换一个服务的同时，所有的服务都将被切换。

```
cluadmin> service relocate
  0) svc01
  c) cancel

Choose service to relocate: 0
Are you sure? (yes/no/?) yes
Relocating svc01. Service  relocated.
```

如果所有的服务添加完成，请直接到第十三章继续阅读随后内容。

第 10 章

配置 LB 服务--TUN 方式

配置服务前要求用户在安装软件和初始化过程中均选择 tun 作为负载均衡技术。GreatTurbo Cluster Server 10 在调度节点正常启动后，就可以开始将用户的应用配置成负载均衡的服务了。本章说明如何配置使用 tun 方式的服务，对服务进行配置的过程如下：

(1)配置 realserver 环境

(2)配置 director 环境

(3)在 director 上配置服务

1. 配置工具

GreatTurbo Cluster Server 10 使用 cluadmin 进行配置服务。

/opt/cluster/bin/cluadmin 是文本界面的管理工具，使用它可以很方便地配置服务。cluadmin 类似 bash，可以使用 TAB 键把命令补全。

注意：如果您使用的是 Turbolinux 中文版，则在运行 cluadmin 之前，请先在 bash 下执行“unset LC_CTYPE”。

2. 配置条件

- 请在 GreatTurbo Cluster Server 10 运行时执行 cluadmin 配置服务。

  注意：GreatTurbo Cluster Server 10 采用时间戳来判断两个调度节点的配置是否同步，所以不要轻易修改系统时间。如果确实需要修改系统时间，请先停止 GreatTurbo Cluster Server 10，修改完系统时间后，再重新启动 GreatTurbo Cluster Server 10。

- 该方式要求所有节点能够加载 ipip 模块。

3. 配置 realserver 环境

1) 以下 2-4 项操作要求在所有的 realserver 上都执行。
2) 根据业务需要修改/etc/init.d/lbrealserver-tun 文件。
   - 如果 realserver 上需要运行的服务没有在机器启动后自动运行,那么需要在上述文件 start、stop 部分的指定位置添加该服务的启动和停止脚本。

✧ 如果 realserver 上运行的服务在机器启动后自动运行，那么不需要修改上述文件。

3) 根据 realserver 上运行的服务个数及配置，修改 /opt/cluster/lb/lbrealserver-conf-tun 文件。

i. 如果 realserver 上运行多个服务，那么需要在 start 部分添加每个服务的浮动 ip 和路由设置，在 stop 部分添加取消路由设置的内容。添加完毕后参考下例对所添加的项目进行修改 如果仅有一个服务，那么只需要修改脚本中已经存在的项既可。

ii. 例，文件中 start 部分有如下项：

```
ifconfig tunl0:0 172.16.70.100 broadcast 172.16.70.100
netmask 255.255.255.255 up
route add -host 172.16.70.100 dev tunl0
```

✧ 将每个服务的浮动 ip 都建立在 tunl0 设备上，分别为 tunl0:0，tunl0:1….。

✧ 修改其中的 ip 地址“172.16.70.100”和 broadcast 地址为服务的实际浮动 ip 地址。

✧ 修改路由设置中的 ip 地址，将“172.16.70.100”修改为服务的真实浮动 ip 地址。

iii. 例，文件中 stop 部分有如下项：

```
route del -host 172.16.70.100 dev tunl0
ifconfig tunl0:0 down
```

✧ 修改路由设置部分的 ip 地址，将“172.16.70.100”修改为服务的真实浮动 ip 地址。

✧ 修改 ifconfig 部分的“tunl0:0”，要求和 start 部分的设置相对应。

4) 执行“/etc/init.d/lbrealserver-tun start”命令，将该节点配置成一个 realserver。如果用户从集群中删掉该节点，那么需要执行“/etc/init.c/lbrealserver-tun stop”命令恢复原来设置。

5) 如果用户在已经运行的 GreatTurbo Load Balance Server 系统上添加服务，那么需要在每个 realserver 上先后进行如下操作

✧ 执行/etc/init.d/lbrealserver-tun stop；

✧ 修改/opt/cluster/lb/lbrealserver-conf-tun 文件；

✧ 执行/etc/init.d/lbrealserver-tun start。

4. 配置调度节点环境

1) 如果调度节点（director）没有作为 realserver 使用，那么不需要进行环境设置。

2) 如果 director 同时作为 realserver 使用，那么需要按照普通 realserver 配置一样首先配置环境。具体配置方法参见“配置 realsever 环境”

5. 配置命令

运行/opt/cluster/bin/cluadmin，然后在 cluadmin 中按下两次 TAB 键，会显示如下所示的所有命令。（[Tab] [Tab]表示连续按两下 Tab 键）

```
cluadmin>[Tab] [Tab]
apropos
clear
exit
help
cluster status
cluster monitor
cluster loglevel
cluster loglevel syncd
cluster loglevel svcmgr
cluster loglevel svccheck
cluster loglevel powerd
cluster loglevel heartbeat
cluster loglevel clumon
cluster heartbeat
cluster mail from
cluster mail to
cluster mail smtpserver
cluster mail level
cluster reload
cluster name
cluster edit
cluster backup
cluster restore
cluster saveas
cluster restorefrom
service add
service show state
service show config
service show services
service modify
service disable
service enable
service relocate
service delete
nbd add
nbd delete
nbd show
help apropos
```

```
help clear
help exit
help help
help cluster status
help cluster monitor
help cluster loglevel
help cluster reload
help cluster name
help cluster edit
help cluster backup
help cluster restore
help cluster saveas
help cluster restorefrom
help service add
help service show state
help service show config
help service show services
help service modify
help service disable
help service enable
help service relocate
help service delete
help nbd add
help nbd delete
help nbd show
```

6. 配置单端口服务

1) 执行 cluadmin

```
[root@test1 root]# cluadmin
Sat May  9 17:48:57 CST 2006
…
```

2) 执行 service add（GreatTurbo Cluster Server 10 最多只支持 256 个服务）

```
cluadmin> service add
…
Currently defined services:
```

3) 输入服务的名字（如果有多个服务，服务的名字不能重复）

```
Service name: svc01
```

4) 输入服务的类别（只能选择 lb）

```
Service type (ha/lb): lb
```

5) 配置服务的优先节点

如果服务有优先节点，当优先节点启动 LB 时或者优先节点的网卡故障恢复时，这个服务将会自动迁移到优先节点上运行。

优先节点缺省值为 None。直接回车表示选择 None，即不选择优先节点。否则输入优先节点的 hostname 名称。

如果配置优先节点，那么所有服务必须选择同一优先节点。

```
Preferred member [None]: test1
```

6) 配置服务的用户脚本

服务的用户脚本是启动和停止应用程序的脚本。配置服务的用户脚本时，请输入全路径名。所有 lb 服务的用户脚本需要保持相同。

如果两个调度节点同时作为 realserver 使用，那么使用 /opt/cluster/bin/lbdirector-real-tun 作为用户脚本。

```
User script (required, e.g., /opt/cluster/bin/lbdirector-dr):
/opt/cluster/bin/lbdirector-tun
```

7) 配置服务检测脚本

所有 lb 服务只能使用同一指定脚本：/opt/cluster/bin/lbcheck

配置服务检测脚本时还需要指定如下的参数：

- “Check interval”是检测服务的时间间隔，建议根据用户应用要求的切换时间设定合适的值，一般输入 5 秒，表示每隔 5 秒时间执行一次检测脚本。
- “Check timeout”是检测脚本执行的超时时间，执行用户检测脚本的超时时间，建议输入 30 秒。
- “Max error count”是允许服务连续错误的次数，一般输入 3 次，表示出错 3 次后，服务进行切换。

```
Do you want to add a check script to the service (yes/no/?) [no]:
yes
```

```
Check Script Information
Check script (e.g., "/opt/cluster/bin/lbcheck" or None) [None]:
/opt/cluster/bin/lbcheck
Check interval (in seconds) [None]: 5
Check timeout (in seconds) [None]: 30
Max error count [None]: 3
```

8) 配置服务的端口号

对于单端口服务只需要输入实际服务运行的端口号即可。如 http 为 80，ssh 为 22。

```
Application port(e.g., 80 or None) [None]: 22
```

9) 配置服务的协议

```
Protocol(tcp/udp) [tcp]: tcp
```

10) 配置服务的浮动 IP

服务可以绑定浮动 IP，浮动 IP 随着服务而浮动，也就是说这个 IP 所在的节点也就是服务所在的节点。

注意：<u>配置浮动 IP 时必须正确输入如下参数，注意随着 IP 地址的不同，相应的子网掩码需与之对应。</u>

- “address”是浮动 IP 的地址。
- “netmask”是浮动 IP 的子网掩码。
- “device” 是浮动 IP 所要绑定的网卡。

```
Virtual IP address(required): 172.16.71.111
Virtual IP network mask(optional) [None]: 255.255.255.0
Virtual IP device(required, e.g., eth0:1): eth0:1
```

11) 配置 firewall mark

单端口服务不需要配置，直接回车即可。

```
Firewall mark [None]:
```

12) 配置真实服务器重新加入集群的时间

建议使用默认值

```
Re-entry time(in seconds) [15]:
```

13) 配置检测 realserver 的时间间隔

建议使用默认值

```
Service timeout(in seconds) [6]:
```

14) 配置服务连接表选项

建议配置为 yes，表示当真实服务器重新加入集群内要把已经存在的连接表清空，避免新加入的服务器过度负载。

```
Quiesce server(yes/no) [no]:yes
```

15) 配置 realserver 的负载监视工具

调度节点能够使用 rup 或 ruptime 来监视各个真正服务器的载量。如果选择了 rup，每个真正服务器就必须运行 rstatd 服务。如果选择了 ruptime，每个真正服务器就必须运行 rwhod 服务。

注意： 载量监视和负载均衡不同。当与加权的调度算法综合使用时，能够导致难以预测的调度行为。还有，如果你要使用载量监视，群集中的真正服务器就必须是 Linux 机器。

```
Load monitoring tool(rup/ruptime/None) [None]:
```

16) 配置调度算法

根据服务的具体要求可以选择一种调度算法。

```
Scheduling(rr/wrr/lc/wlc/lblc/lblcr/dh/sh) [wlc]:
```

17) 配置持续连接时间

单端口服务不需要配置此项，回车即可。

```
Persistence(in seconds) [None]:
```

18) 配置 realserver 上服务的监视脚本

这里输入监视 realserver 上服务的脚本。我们提供了一些监视脚本，可以用来监视大多数常见服务；如果没有提供指定服务的脚本，您可以参照其他监视脚本自己编写或者选择 None。如果选择使用脚本检测 realserver，注意要在脚本路径后面使用特殊的“%h”符号。调度节点的 Nanny 进程会在调度该脚本时，使用 realserver 的真实 ip 地址代替%h。

```
Sending program(e.g., "/opt/cluster/bin/lbsend-http %h" or None) [None]: /opt/cluster/bin/lbsend-ssh %h
```

19) 配置 realserver 的回复文本

如果使用我们提供的脚本，那么要输入“OK”；如果使用其它方式，那么 expect 项根据具体脚本情况填写。

```
Expect string(e.g., "OK" or None) [None]: OK
```

20) 配置回复文本属性

如果回复文本时正则表达式，则此处选择 yes；否则选择 no

```
Treat expect string as a regular expression(yes/no) [no]: no
```

21) 配置 realserver

一个服务可以配置多个 realserver，调度器根据调度算法，将客户的请求调度到不同的 realserver 上。每个 realserver 都需要配置如下几项：

- “Server name”是一个用于真正服务器的描述性名称。
- “Server address”是真正服务器的 IP 地址。
- “Server weight” 是一个表明和集群内其它主机相比而言的主机能力的整数值。
- “Server active” 表示真实服务器是否启用。

```
Do you want to add server to the service (yes/no/?) [no]: y

        Server Information

Server name(required): server1
```

```
Server address(required): 172.16.70.211
Server weight [None]: 1
Server active(yes/no) [yes]: yes
Do you want to (a)dd, (m)odify, (d)elete or (s)how a server, or
are you (f)inished adding servers: f
```

22) 设置停止服务失败时是否 reboot 机器

如果设为“ yes ”，当服务停止失败时，为了释放服务的资源，将自动 reboot 机器。

如果设为“ no ”，当服务停止失败时，不会自动 reboot 机器，需要用户手动干预。

```
Reboot system if stop the service failed (yes/no/?) [yes]: yes
```

23) 设置是否 disable 服务

如果设为“ yes ”，服务将不会立即被启动，只能以后由用户手动启动。

如果设为“ no ”，服务将立即被启动。一般选择 no，表示立即启动服务。

```
Disable service (yes/no/?) [no]: no
```

24) 保存服务配置

选择 yes

```
Add test service as shown? (yes/no/?) yes
```

到这里，我们成功的添加了一个单端口的服务。如果用户还需要添加服务，那么请重复上述过程。

7. 修改服务

如果需要修改服务的配置，请在 cluadmin 下执行 service modify 命令。

首先要求选择需要修改的服务，请输入列表中的服务的编号。

如果服务正在运行，将要求先停止服务，请输入 yes。

然后所有的操作都和添加服务相同。

```
cluadmin> service modify
```

```
  You will be prompted for information about the service.

  Enter a question mark (?) at a prompt to obtain help.

  Enter a colon (:) and a single-character command at a prompt to
do
  one of the following:

  c - Cancel and return to the top-level cluadmin command
  r - Restart to the initial prompt while keeping previous responses
  p - Proceed with the next prompt



  0) svc01
  c) cancel

Choose service to modify: 0
Modifying: svc01
Service is not disabled.  Disable it? (yes/no/?) yes
```

8. 删除服务

如果要删除服务，请执行 cluadmin 的 service delete 命令。

然后输入列表中的服务的编号即可。

```
cluadmin> service delete
  0) svc01
  c) cancel

Choose service to delete: 0
Deleting svc01, are you sure? (yes/no/?): yes
Service svc01 disabled
Svc01 deleted.
```

9. 启动服务

如果在添加服务时没有启动服务，需要手动启动服务，请执行 cluadmin 的 service enable 命令。

首先要求选择需要启动的服务，请输入列表中的服务的编号。

然后要求选择在哪个节点上启动服务，请输入节点的编号。

注意：启动一个服务的同时，所有的服务都将被启动。

```
cluadmin> service enable
  0) svc01
  c) cancel

Choose service to enable: 0
Are you sure? (yes/no/?) yes
  0) test1
  1) test2
  c) cancel

Choose member: 0
Enabling svc01 on member test1. Service  enabled.
```

10. 停止服务

如果需要停止服务，请执行 cluadmin 的 service disable 命令。

首先选择要停止的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

注意： 停止一个服务的同时，所有的服务都将被停止。

```
cluadmin> service disable
  0) svc01
  c) cancel

Choose service to disable: 0
Are you sure? (yes/no/?) yes
Disabling svc01. Service  disabled.
```

11. 切换服务

如果需要把在一个节点上运行的服务切换到另一个节点上，请执行 cluadmin 的 service relocate 命令。

首先选择要切换的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

注意：切换一个服务的同时，所有的服务都将被切换。

```
cluadmin> service relocate
```

```
  0) svc01
  c) cancel

Choose service to relocate: 0
Are you sure? (yes/no/?) yes
Relocating svc01. Service  relocated.
```

如果所有的服务添加完成，请直接到第十三章继续阅读随后内容。

第 11 章

配置 LB 服务--NAT 方式

配置服务前要求用户在安装软件和初始化过程中均选择 nat 作为负载均衡技术。

GreatTurbo Cluster Server 10 在调度节点正常启动后，就可以开始将用户的应用配置成负载均衡的服务了。本章说明如何配置使用 nat 方式的服务，对服务进行配置的过程如下：

(1)配置 realserver 环境

(2)配置 director 环境

(3)在 director 上配置服务

1. 配置工具

GreatTurbo Cluster Server 10 使用 cluadmin 进行配置服务。

/opt/cluster/bin/cluadmin 是文本界面的管理工具，使用它可以很方便地配置服务。cluadmin 类似 bash，可以使用 TAB 键进行命令补全。

注意：如果您使用的是 Turbolinux 中文版，则在运行 cluadmin 之前，请先在 bash 下执行“unset LC_CTYPE”。

2. 配置条件

- 请在 GreatTurbo Cluster Server 10 运行时执行 cluadmin 配置服务。

  注意：GreatTurbo Cluster Server 10 采用时间戳来判断两个调度节点的配置是否同步，所以不要轻易修改系统时间。如果确实需要修改系统时间，请先停止 GreatTurbo Cluster Server 10，修改完系统时间后，再重新启动 GreatTurbo Cluster Server 104。

- NAT 方式要求普通的 realserver 不能有提供对外访问的网卡，所有网络设备都要设置为私有 IP，并且需要在同一个网段内。

3. 配置 realserver 环境

1） 以下 2-4 步操作要求在所有的 realserver 上都执行。
2） 根据业务需要修改/etc/init.d/lbrealserver-nat 文件。
    - 如果 realserver 上需要运行的服务没有在机器启动后自动运行，那么需要在上述文件 start、stop 部分的指定位置添加该服务的启动和停止脚本。

- ✧ 如果 realserver 上运行的服务在机器启动后自动运行,那么不需要修改上述文件。

3) 修改/opt/cluster/lb/lbrealserver-conf-nat 文件。
   i. 修改 start 和 stop 部分的网关地址。Start 部分将默认网关改为 nat router 地址；stop 部分将默认网关恢复为原来的网关。
   ii. 例，文件中 start 部分有如下项：
      route add default gw 192.168.0.3
      ✧ 将网关地址“192.168.0.3”改为 nat router 地址。
   iii. stop 部分有如下项：
      route add default gw 172.16.68.1
      ✧ 将网关地址“172.16.68.1”改为原来的网关地址。
4) 执行“/etc/init.d/lbrealserver-nat start”命令，将该节点配置成一个 realserver。如果用户从集群中删掉该节点，那么需要执行“/etc/init.c/lbrealserver-nat stop”命令恢复原来设置。
5) 如果用户在已经运行的 GreatTurbo Load Balance Server 系统上添加服务，那么需要在每个 realserver 上先后进行如下操作
   ✧ 执行/etc/init.d/lbrealserver-nat stop；
   ✧ 修改/opt/cluster/lb/lbrealserver-conf-nat 文件；
   ✧ 执行/etc/init.d/lbrealserver-nat start。

4. 配置调度节点环境

1) 如果调度节点（director）没有作为 realserver 使用，那么 director 两节点不需要其他配置。
2) 如果 director 同时作为 realserver 使用，那么需要对/opt/cluster/lb/lbrealserver-conf-nat 文件进行额外的修改。
   i. 假如两调度节点的 eth0 网卡提供对外访问，该网卡的 ip 地址为“172.16.70.100”，那么要在 start 部分添加
      route del –net 172.16.70.0 netmask 255.255.252.0 dev eth0
      要保证所有的 eth0 路由全部被删除。
   ii. 在 stop 部分添加
      route add –net 172.16.70.0 netmask 255.255.252.0 gw 172.16.68.1 dev eth0
      确保恢复到机器原来的路由表。
   iii. 路由增删的具体配置方法请咨询网络管理员。
   iv. 因为删掉对外访问路由后会使对外的网络访问失效，如果该网卡配置为 GreatTurbo Cluster Server 10 的心跳，那么该心跳就会失效。为了解决上述问题，我们还需要加上一个新路由，仅用来提供“172.16.70.100”ip 地址的路由：
      route add –host 172.16.70.100 dev eth0
      这样配置后，心跳线也不会受到影响。

5. 配置命令

运行/opt/cluster/bin/cluadmin，然后在 cluadmin 中按下两次 TAB 键，会显示如下所示的所有命令。（[Tab] [Tab]表示连续按两下 Tab 键）

```
cluadmin>[Tab] [Tab]
apropos
clear
exit
help
cluster status
cluster monitor
cluster loglevel
cluster loglevel syncd
cluster loglevel svcmgr
cluster loglevel svccheck
cluster loglevel powerd
cluster loglevel heartbeat
cluster loglevel clumon
cluster heartbeat
cluster mail from
cluster mail to
cluster mail smtpserver
cluster mail level
cluster reload
cluster name
cluster edit
cluster backup
cluster restore
cluster saveas
cluster restorefrom
service add
service show state
service show config
service show services
service modify
service disable
service enable
service relocate
service delete
nbd add
nbd delete
nbd show
help apropos
```

```
help clear
help exit
help help
help cluster status
help cluster monitor
help cluster loglevel
help cluster reload
help cluster name
help cluster edit
help cluster backup
help cluster restore
help cluster saveas
help cluster restorefrom
help service add
help service show state
help service show config
help service show services
help service modify
help service disable
help service enable
help service relocate
help service delete
help nbd add
help nbd delete
help nbd show
```

6. 配置单端口服务

1) 执行 cluadmin

```
[root@test1 root]# cluadmin
Sat May  9 17:48:57 CST 2006
…
```

2) 执行 service add（GreatTurbo Cluster Server 10 最多只支持 256 个服务）

```
cluadmin> service add
…
Currently defined services:
```

3) 输入服务的名字（如果有多个服务，服务的名字不能重复）

```
Service name: svc01
```

4) 输入服务的类别（只能选择 lb）

```
Service type (ha/lb): lb
```

5) 配置服务的优先节点

如果服务有优先节点，当优先节点启动 LB 时或者优先节点的网卡故障恢复时，这个服务将会自动迁移到优先节点上运行。

优先节点缺省值为 None。直接回车表示选择 None，即不选择优先节点。否则输入优先节点的 hostname 名称。

如果配置优先节点，那么所有服务必须选择同一优先节点。

```
Preferred member [None]: test1
```

6) 配置服务的用户脚本

服务的用户脚本是启动和停止应用程序的脚本。配置服务的用户脚本时，请输入全路径名。所有 lb 服务的用户脚本需要保持相同。

如果两个调度节点同时作为 realserver 使用，那么使用 /opt/cluster/bin/lbdirector-real-nat 作为用户脚本。

```
User script (required, e.g., /opt/cluster/bin/lbdirector-dr):
/opt/cluster/bin/lbdirector-nat
```

7) 配置服务检测脚本

所有 lb 服务只能使用同一指定脚本：/opt/cluster/bin/lbcheck

配置服务检测脚本时还需要指定如下的参数：

- "Check interval"是检测服务的时间间隔，建议根据用户应用要求的切换时间设定合适的值，一般输入 5 秒，表示每隔 5 秒时间执行一次检测脚本。
- "Check timeout"是检测脚本执行的超时时间，执行用户检测脚本的超时时间，建议输入 30 秒。
- "Max error count"是允许服务连续错误的次数，一般输入 3 次，表示出错 3 次后，服务进行切换。

```
Do you want to add a check script to the service (yes/no/?) [no]:
yes
```

```
Check Script Information
Check script (e.g., "/opt/cluster/bin/lbcheck" or None) [None]:
/opt/cluster/bin/lbcheck
Check interval (in seconds) [None]: 5
Check timeout (in seconds) [None]: 30
Max error count [None]: 3
```

8) 配置服务的端口号

对于单端口服务只需要输入实际服务运行的端口号即可。如 http 为 80，ssh 为 22。

```
Application port(e.g., 80 or None) [None]: 22
```

9) 配置服务的协议

```
Protocol(tcp/udp) [tcp]: tcp
```

10) 配置服务的浮动 IP

服务可以绑定浮动 IP，浮动 IP 随着服务而浮动，也就是说这个 IP 所在的节点也就是服务所在的节点。

注意：<u>配置浮动 IP 时必须正确输入如下参数，注意随着 IP 地址的不同，相应的子网掩码需与之对应。</u>

- “address”是浮动 IP 的地址。
- “netmask”是浮动 IP 的子网掩码。
- “device” 是浮动 IP 所要绑定的网卡。

```
Virtual IP address(required): 172.16.71.111
Virtual IP network mask(optional) [None]: 255.255.255.0
Virtual IP device(required, e.g., eth0:1): eth0:1
```

11) 配置 firewall mark

单端口服务不需要配置，直接回车即可。

```
Firewall mark [None]:
```

12) 配置真实服务器重新加入集群的时间

建议使用默认值

```
Re-entry time(in seconds) [15]:
```

13) 配置检测 realserver 的时间间隔

建议使用默认值

```
Service timeout(in seconds) [6]:
```

14) 配置服务连接表选项

建议配置为 yes，表示当真实服务器重新加入集群内要把已经存在的连接表清空，避免新加入的服务器过度负载。

```
Quiesce server(yes/no) [no]:yes
```

15) 配置 realserver 的负载监视工具

调度节点能够使用 rup 或 ruptime 来监视各个真正服务器的载量。如果选择了 rup，每个真正服务器就必须运行 rstatd 服务。如果选择了 ruptime，每个真正服务器就必须运行 rwhod 服务。

注意： 载量监视和负载均衡不同。当与加权的调度算法综合使用时，能够导致难以预测的调度行为。还有，如果你要使用载量监视，群集中的真正服务器就必须是 Linux 机器。

```
Load monitoring tool(rup/ruptime/None) [None]:
```

16) 配置调度算法

根据服务的具体要求可以选择一种调度算法。

```
Scheduling(rr/wrr/lc/wlc/lblc/lblcr/dh/sh) [wlc]:
```

17) 配置持续连接时间

单端口服务不需要配置此项，回车即可。

```
Persistence(in seconds) [None]:
```

18) 配置 realserver 上服务的监视脚本

这里输入监视 realserver 上服务的脚本。我们提供了一些监视脚本，可以用来监视大多数常见服务；如果没有提供指定服务的脚本，您可以参照其他监视脚本自己编写或者选择 None。如果选择使用脚本检测 realserver，注意要在脚本路径后面使用特殊的“%h”符号。调度节点的 Nanny 进程会在调度该脚本时，使用 realserver 的真实 ip 地址代替%h。

```
Sending program(e.g., "/opt/cluster/bin/lbsend-http %h" or
None) [None]: /opt/cluster/bin/lbsend-ssh %h
```

19) 配置 realserver 的回复文本

如果使用我们提供的脚本，那么要输入“OK”；如果使用其它方式，那么 expect 项根据具体脚本情况填写。

```
Expect string(e.g., "OK" or None) [None]: OK
```

20) 配置回复文本属性

如果回复文本时正则表达式，则此处选择 yes；否则选择 no

```
Treat expect string as a regular expression(yes/no) [no]: no
```

21) 配置 realserver

一个服务可以配置多个 realserver，调度器根据调度算法，将客户的请求调度到不同的 realserver 上。每个 realserver 都需要配置如下几项：

- “Server name”是一个用于真正服务器的描述性名称。
- “Server address”是真正服务器的 IP 地址。
- “Server weight” 是一个表明和集群内其它主机相比而言的主机能力的整数值。
- “Server active” 表示真实服务器是否启用。

```
Do you want to add server to the service (yes/no/?) [no]: y

        Server Information

Server name(required): server1
```

```
Server address(required): 192.168.0.18
Server weight [None]: 1
Server active(yes/no) [yes]: yes
Do you want to (a)dd, (m)odify, (d)elete or (s)how a server, or
are you (f)inished adding servers: f
```

22) 设置停止服务失败时是否 reboot 机器

如果设为“ yes ”，当服务停止失败时，为了释放服务的资源，将自动 reboot 机器。

如果设为“ no ”，当服务停止失败时，不会自动 reboot 机器，需要用户手动干预。

```
Reboot system if stop the service failed (yes/no/?) [yes]: yes
```

23) 设置是否 disable 服务

如果设为“ yes ”，服务将不会立即被启动，只能以后由用户手动启动。

如果设为“ no ”，服务将立即被启动。一般选择 no，表示立即启动服务。

```
Disable service (yes/no/?) [no]: no
```

24) 保存服务配置

选择 yes

```
Add test service as shown? (yes/no/?) yes
```

到这里，我们成功的添加了一个单端口的服务。如果用户还需要添加服务，请重复上述过程。

7. 修改服务

如果需要修改服务的配置，请在 cluadmin 下执行 service modify 命令。

首先要求选择需要修改的服务，请输入列表中的服务的编号。

如果服务正在运行，将要求先停止服务，请输入 yes。

然后所有的操作都和添加服务相同。

```
cluadmin> service modify
```

```
  You will be prompted for information about the service.

  Enter a question mark (?) at a prompt to obtain help.

  Enter a colon (:) and a single-character command at a prompt to
do
  one of the following:

  c - Cancel and return to the top-level cluadmin command
  r - Restart to the initial prompt while keeping previous responses
  p - Proceed with the next prompt



  0) svc01
  c) cancel

Choose service to modify: 0
Modifying: svc01
Service is not disabled.  Disable it? (yes/no/?) yes
```

8. 删除服务

如果要删除服务，请执行 cluadmin 的 service delete 命令。

然后输入列表中的服务的编号即可。

```
cluadmin> service delete
  0) svc01
  c) cancel

Choose service to delete: 0
Deleting svc01, are you sure? (yes/no/?): yes
Service svc01 disabled
Svc01 deleted.
```

9. 启动服务

如果在添加服务时没有启动服务，需要手动启动服务，请执行 cluadmin 的 service enable 命令。

首先要求选择需要启动的服务，请输入列表中的服务的编号。

然后要求选择在哪个节点上启动服务，请输入节点的编号。

注意：启动一个服务的同时，所有的服务都将被启动。

```
cluadmin> service enable
  0) svc01
  c) cancel

Choose service to enable: 0
Are you sure? (yes/no/?) yes
  0) test1
  1) test2
  c) cancel

Choose member: 0
Enabling svc01 on member test1. Service  enabled.
```

10. 停止服务

如果需要停止服务，请执行 cluadmin 的 service disable 命令。

首先选择要停止的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

注意： 停止一个服务的同时，所有的服务都将被停止。

```
cluadmin> service disable
  0) svc01
  c) cancel

Choose service to disable: 0
Are you sure? (yes/no/?) yes
Disabling svc01. Service  disabled.
```

11. 切换服务

如果需要把在一个节点上运行的服务切换到另一个节点上，请执行 cluadmin 的 service relocate 命令。

首先选择要切换的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

注意：切换一个服务的同时，所有的服务都将被切换。

```
cluadmin> service relocate
```

```
  0) svc01
  c) cancel

Choose service to relocate: 0
Are you sure? (yes/no/?) yes
Relocating svc01. Service  relocated.
```

如果所有的服务添加完成，请到第十三章继续阅读随后内容。

第 12 章

配置 LB 服务--多端口 DR 方式

配置服务前要求用户在安装软件和初始化过程中均选择 dr 作为负载均衡技术。GreatTurbo Cluster Server 10 在调度节点正常启动后，就可以开始将用户的应用配置成负载均衡的服务了。本章说明如何配置使用 dr 方式的多端口服务，对服务进行配置的过程如下：

1) 配置 realserver 环境
2) 配置 director 环境
3) 在 director 上配置多端口服务

1. 配置工具

GreatTurbo Cluster Server 10 使用 cluadmin 进行配置服务。

/opt/cluster/bin/cluadmin 是文本界面的管理工具，使用它可以很方便地配置服务。cluadmin 类似 bash，可以使用 TAB 键进行命令补全。

注意：如果您使用的是 Turbolinux 中文版，则在运行 cluadmin 之前，请先在 bash 下执行“unset LC_CTYPE”。

2. 配置条件

- 请在 GreatTurbo Cluster Server 10 运行时执行 cluadmin 配置服务。

  注意：GreatTurbo Cluster Server 10 采用时间戳来判断两个调度节点的配置是否同步，所以不要轻易修改系统时间。如果确实需要修改系统时间，请先停止 GreatTurbo Cluster Server 10，修改完系统时间后，再重新启动 GreatTurbo Cluster Server 10。
- 因为 dr 工作方式需要使用防火墙（iptables）作为转发工具，所以需要保证在作 realserver 的机器上安装 iptables 软件。
- 调度节点和所有 realserver 都必须在物理上有一个网卡通过不分断的局域网相连，如通过高速的交换机或者 HUB 相连。

3. 配置 realserver 环境

1) 以下 2-4 步操作要求在所有的 realserver 上都执行。
2) 根据业务需要修改/etc/init.d/lbrealserver-dr 文件。

- ✧ 如果 realserver 上需要运行的服务没有在机器启动后自动运行，那么需要在上述文件 start、stop 部分的指定位置添加该服务的启动和停止脚本。
- ✧ 如果 realserver 上运行的服务在机器启动后自动运行，那么不需要修改上述文件。

3) 根据 realserver 上运行的服务个数及配置，修改 /opt/cluster/lb/lbrealserver-conf-dr 文件。
   i. 如果 realserver 上运行 n 个服务，那么需要在 start、stop 部分复制 n-1 个 iptables 项。然后根据服务具体情况修改每一项的 ip 地址、dport、--to-port 值。如果仅有一个服务，那么只需要修改脚本中已经存在的 iptalbes 项。
   ii. 例，文件中 start 部分有如下项：
   ```
   iptables -t nat -A PREROUTING -p tcp -d 172.16.70.100
   --dport 22 -j REDIRECT --to-port 22
   ```
   - ✧ 修改其中的 ip 地址“172.16.70.100”为服务的实际浮动 ip 地址。
   - ✧ 把 dport 后和--to-port 后的端口号“22”修改为服务实际占用的端口。
   - ✧ 根据上述内容，在 stop 部分作相应的修改，清除所设置的防火墙规则。

4) 执行“/etc/init.d/lbrealserver-dr start”命令，将该节点配置成一个 realserver。如果用户从集群中删掉该节点，那么需要执行“/etc/init.c/lbrealserver-dr stop”命令恢复原来设置。
5) 如果用户在已经运行的 GreatTurbo Load Balance Server 系统上添加服务，那么需要在每个 realserver 上先后进行如下操作。
   - ✧ 执行/etc/init.d/lbrealserver-dr stop；
   - ✧ 修改/opt/cluster/lb/lbrealserver-conf-dr 文件；
   - ✧ 执行/etc/init.d/lbrealserver-dr start

4. 配置调度节点环境

1) 多端口服务需要在 director 上建立一些防火墙标记，来使 director 可以正确处理多端口服务。修改/opt/cluster/bin/lbdirector-conf-fwm 脚本，在 start 部分和 stop 部分根据服务的数量和服务所使用的端口号来做相应修改，设置防火墙标记。
2) start 部分有如下项：
   ```
   iptables -t mangle -A PREROUTING -p tcp -d 172.16.70.150/32
   --dport 21 -j MARK --set-mark 21
   ```
   - ✧ 如果服务占用 n 个端口，那么上述内容需要有 n 份。
   - ✧ 把浮动 ip 地址“172.16.70.150”修改为服务真正的浮动 ip 地址。
   - ✧ 把端口号（dport 部分）修改为服务所使用的一个端口。
   - ✧ 最后设定 firemark 的值，也就是—set-mark 部分。这里使用的是 21，用户可以根据需要另外选定其它值。Firemark 的值在添加 LB 服务的时候需要使用。
3) stop 部分要根据 start 部分作相应的修改。

4) 如果用户有 n 个服务，那么需要重复上述过程 n 次。

5. 配置命令

运行/opt/cluster/bin/cluadmin，然后在 cluadmin 中按下两次 TAB 键，会显示如下所示的所有命令。([Tab] [Tab]表示连续按两下 Tab 键)

```
cluadmin>[Tab] [Tab]
apropos
clear
exit
help
cluster status
cluster monitor
cluster loglevel
cluster loglevel syncd
cluster loglevel svcmgr
cluster loglevel svccheck
cluster loglevel powerd
cluster loglevel heartbeat
cluster loglevel clumon
cluster heartbeat
cluster mail from
cluster mail to
cluster mail smtpserver
cluster mail level
cluster reload
cluster name
cluster edit
cluster backup
cluster restore
cluster saveas
cluster restorefrom
service add
service show state
service show config
service show services
service modify
service disable
service enable
service relocate
service delete
nbd add
nbd delete
```

```
nbd show
help apropos
help clear
help exit
help help
help cluster status
help cluster monitor
help cluster loglevel
help cluster reload
help cluster name
help cluster edit
help cluster backup
help cluster restore
help cluster saveas
help cluster restorefrom
help service add
help service show state
help service show config
help service show services
help service modify
help service disable
help service enable
help service relocate
help service delete
help nbd add
help nbd delete
help nbd show
```

6. 配置多端口服务

1) 执行 cluadmin

```
[root@test1 root]# cluadmin
Sat May  9 17:48:57 CST 2006
…
```

2) 执行 service add（GreatTurbo Cluster Server 10 最多只支持 256 个服务）

```
cluadmin> service add
…
Currently defined services:
```

3) 输入服务的名字（如果有多个服务，服务的名字不能重复）

```
Service name: svc01
```

4) 输入服务的类别（只能选择 lb）

```
Service type (ha/lb): lb
```

5) 配置服务的优先节点

如果服务有优先节点，当优先节点启动 LB 时或者优先节点的网卡故障恢复时，这个服务将会自动迁移到优先节点上运行。

优先节点缺省值为 None。直接回车表示选择 None，即不选择优先节点。否则输入优先节点的 hostname 名称。

如果配置优先节点，那么所有服务必须选择同一优先节点。

```
Preferred member [None]: test1
```

6) 配置服务的用户脚本

服务的用户脚本是启动和停止应用程序的脚本。配置服务的用户脚本时，请输入全路径名。所有 lb 服务的用户脚本需要保持相同。

如果两个调度节点同时作为 realserver 使用，那么使用 /opt/cluster/bin/lbdirector-real-fwm-dr 作为用户脚本。

```
User script (required, e.g., /opt/cluster/bin/lbdirector-dr):
/opt/cluster/bin/lbdirector-fwm-dr
```

7) 配置服务检测脚本

所有 lb 服务只能使用同一指定脚本：/opt/cluster/bin/lbcheck

配置服务检测脚本时还需要指定如下的参数：

- “Check interval”是检测服务的时间间隔，建议根据用户应用要求的切换时间设定合适的值，一般输入 5 秒，表示每隔 5 秒时间执行一次检测脚本。
- “Check timeout”是检测脚本执行的超时时间，执行用户检测脚本的超时时间，建议输入 30 秒。
- “Max error count”是允许服务连续错误的次数，一般输入 3 次，表示出错 3 次后，服务进行切换。

```
Do you want to add a check script to the service (yes/no/?) [no]:
yes
Check Script Information
Check script (e.g., "/opt/cluster/bin/lbcheck" or None) [None]:
/opt/cluster/bin/lbcheck
Check interval (in seconds) [None]: 5
Check timeout (in seconds) [None]: 30
Max error count [None]: 3
```

8) 配置服务的端口号

对于多端口服务该项可以不用填写。

```
Application port(e.g., 80 or None) [None]:
```

9) 配置服务的协议

```
Protocol(tcp/udp) [tcp]: tcp
```

10) 配置服务的浮动 IP

服务可以绑定浮动 IP，浮动 IP 随着服务而浮动，也就是说这个 IP 所在的节点也就是服务所在的节点。

注意：配置浮动 IP 时必须正确输入如下参数，注意随着 IP 地址的不同，相应的子网掩码需与之对应。

- “address”是浮动 IP 的地址。
- “netmask”是浮动 IP 的子网掩码。
- “device” 是浮动 IP 所要绑定的网卡。

```
Virtual IP address(required): 172.16.71.111
Virtual IP network mask(optional) [None]: 255.255.255.0
Virtual IP device(required, e.g., eth0:1): eth0:1
```

11) 配置 firewall mark

输入/opt/cluster/bin/lbdirector-conf-fwm 文件中对该服务所设置的 firewall 标记值。

```
Firewall mark [None]: 21
```

12) 配置真实服务器重新加入集群的时间

建议使用默认值

```
Re-entry time(in seconds) [15]:
```

13) 配置检测 realserver 的时间间隔

建议使用默认值

```
Service timeout(in seconds) [6]:
```

14) 配置服务连接表选项

建议配置为 yes，表示当真实服务器重新加入集群内要把已经存在的连接表清空，避免新加入的服务器过度负载。

```
Quiesce server(yes/no) [no]:yes
```

15) 配置 realserver 的负载监视工具

调度节点能够使用 rup 或 ruptime 来监视各个真正服务器的载量。如果选择了 rup，每个真正服务器就必须运行 rstatd 服务。如果选择了 ruptime，每个真正服务器就必须运行 rwhod 服务。

注意： 载量监视和负载均衡不同。当与加权的调度算法综合使用时，能够导致难以预测的调度行为。还有，如果你要使用载量监视，群集中的真正服务器就必须是 Linux 机器。

```
Load monitoring tool(rup/ruptime/None) [None]:
```

16) 配置调度算法

根据服务的具体要求可以选择一种调度算法。

```
Scheduling(rr/wrr/lc/wlc/lblc/lblcr/dh/sh) [wlc]:
```

17) 配置持续连接时间

根据服务的具体情况填写持续连接时间。在此时间内由同一个 ip 发来的请求包将被调度到同一个 realserver。

```
Persistence(in seconds) [None]:
```

18) 配置 realserver 上服务的监视脚本

这里输入监视 realserver 上服务的脚本。我们提供了一些监视脚本，可以用来监视大多数常见服务；如果没有提供指定服务的脚本，您可以参照其他监视脚本自己编写或者选择 None。如果选择使用脚本检测 realserver，注意要在脚本路径后面使用特殊的“%h”符号。调度节点的 Nanny 进程会在调度该脚本时，使用 realserver 的真实 ip 地址代替%h。

```
Sending  program(e.g.,  "/opt/cluster/bin/lbsend-http  %h"  or
None) [None]: /opt/cluster/bin/lbsend-ssh %h
```

19) 配置 realserver 的回复文本

如果使用我们提供的脚本，那么要输入“OK”；如果使用其它方式，那么 expect 项根据具体脚本情况填写。

```
Expect string(e.g., "OK" or None) [None]: OK
```

20) 配置回复文本属性

如果回复文本时正则表达式，则此处选择 yes；否则选择 no

```
Treat expect string as a regular expression(yes/no) [no]: no
```

21) 配置 realserver

一个服务可以配置多个 realserver，调度器根据调度算法，将客户的请求调度到不同的 realserver 上。每个 realserver 都需要配置如下几项：

- “Server name”是一个用于真正服务器的描述性名称。
- “Server address”是真正服务器的 IP 地址。
- “Server weight” 是一个表明和集群内其它主机相比而言的主机能力的整数值。
- “Server active” 表示真实服务器是否启用。

```
Do you want to add server to the service (yes/no/?) [no]: y

        Server Information
```

```
Server name(required): server1
Server address(required): 172.16.70.211
Server weight [None]: 1
Server active(yes/no) [yes]: yes
Do you want to (a)dd, (m)odify, (d)elete or (s)how a server, or
are you (f)inished adding servers: f
```

22) 设置停止服务失败时是否 reboot 机器

如果设为“yes”，当服务停止失败时，为了释放服务的资源，将自动 reboot 机器。

如果设为“no”，当服务停止失败时，不会自动 reboot 机器，需要用户手动干预。

```
Reboot system if stop the service failed (yes/no/?) [yes]: yes
```

23) 设置是否 disable 服务

如果设为“yes”，服务将不会立即被启动，只能以后由用户手动启动。

如果设为“no”，服务将立即被启动。一般选择 no，表示立即启动服务。

```
Disable service (yes/no/?) [no]: no
```

24) 保存服务配置

选择 yes

```
Add test service as shown? (yes/no/?) yes
```

到这里，我们成功的添加了一个多端口的服务。如果用户还需要添加服务，那么重复上述过程。

7. 修改服务

如果需要修改服务的配置，请在 cluadmin 下执行 service modify 命令。

首先要求选择需要修改的服务，请输入列表中的服务的编号。

如果服务正在运行，将要求先停止服务，请输入 yes。

然后所有的操作都和添加服务相同。

```
cluadmin> service modify

  You will be prompted for information about the service.

  Enter a question mark (?) at a prompt to obtain help.

  Enter a colon (:) and a single-character command at a prompt to
do
  one of the following:

  c - Cancel and return to the top-level cluadmin command
  r - Restart to the initial prompt while keeping previous responses
  p - Proceed with the next prompt



  0) svc01
  c) cancel

Choose service to modify: 0
Modifying: svc01
Service is not disabled.  Disable it? (yes/no/?) yes
```

8. 删除服务

如果要删除服务，请执行 cluadmin 的 service delete 命令。

然后输入列表中的服务的编号即可。

```
cluadmin> service delete
  0) svc01
  c) cancel

Choose service to delete: 0
Deleting svc01, are you sure? (yes/no/?): yes
Service svc01 disabled
Svc01 deleted.
```

9. 启动服务

如果在添加服务时没有启动服务，需要手动启动服务，请执行 cluadmin 的 service enable 命令。

首先要求选择需要启动的服务，请输入列表中的服务的编号。

然后要求选择在哪个节点上启动服务，请输入节点的编号。

注意：启动一个服务的同时，所有的服务都将被启动。

```
cluadmin> service enable
  0) svc01
  c) cancel

Choose service to enable: 0
Are you sure? (yes/no/?) yes
  0) test1
  1) test2
  c) cancel

Choose member: 0
Enabling svc01 on member test1. Service  enabled.
```

10. 停止服务

如果需要停止服务，请执行 cluadmin 的 service disable 命令。

首先选择要停止的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

注意： 停止一个服务的同时，所有的服务都将被停止。

```
cluadmin> service disable
  0) svc01
  c) cancel

Choose service to disable: 0
Are you sure? (yes/no/?) yes
Disabling svc01. Service  disabled.
```

11. 切换服务

如果需要把在一个节点上运行的服务切换到另一个节点上，请执行 cluadmin 的 service relocate 命令。

首先选择要切换的服务，请输入列表中的服务的编号。

然后要求确认，请输入 yes。

注意：切换一个服务的同时，所有的服务都将被切换。

```
cluadmin> service relocate
  0) svc01
  c) cancel

Choose service to relocate: 0
Are you sure? (yes/no/?) yes
Relocating svc01. Service  relocated.
```

如果所有的服务添加完成，请直接到第十三章继续阅读随后内容。

第 13 章

使用 guiadmin 来配置 GreatTurbo Cluster Server 10

GreatTurbo Cluster Server 10 在两个节点均正常启动后，就可以开始将用户的应用配置成服务了。用户可以使用 cluadmin 来进行配置，也可以使用更方便、快捷的图形配置工具 guiadmin 来进行配置。guiadmin 可以跨平台、远程的配置服务，还具有配置集群信息，监控 GreatTurbo Cluster Server 10 的状态等一些功能。

Guiadmin 目前只支持对 HA 服务的管理，不支持对 LB 服务的管理

本章将介绍 guiadmin 的基本功能，并说明如何使用 guiadmin 来对 GreatTurbo Cluster Server 10 进行配置和监控。

guiadmin 的基本功能如下：

1) 配置集群的公共信息
2) 配置高可用的 HA 服务
3) 对指定服务进行状态变更
4) 修改和删除指定 HA 服务信息
5) 显示两节点的配置信息
6) 实时监控 GreatTurbo Cluster Server 10 状态信息

1. guiadmin 说明

guiadmin 是使用 java 技术实现的跨平台的图形配置工具，由客户端和服务器端两部分组成，均采用 java 技术实现。客户端和服务器通过网络传送数据，实现对 GreatTurbo Cluster Server 10 配置信息的更改。服务器端软件要在 GreatTurbo Cluster Server 10 的两个节点上安装，客户端软件可以安装在两个节点上，也可以安装在其他任意平台，如：windows，linux 等。C-S 结构的设计可以使用户可以对 GreatTurbo Cluster Server 10 系统进行远程控制。

当使用配置工具的时候，必须保证 guiadmin 服务器端进程已经在 GreatTurbo Cluster Server 10 的节点启动，客户端进程启动后需要主动去和已经启动了服务的节点进行连接。

用户对 GreatTurbo Cluster Server 10 的所有操作只需要在客户端进行，guiadmin 会自动同步配置两个节点之间的信息。

guiadmin 也可以实时的获取 GreatTurbo Cluster Server 10 状态信息，实现远程监控。

注意：必须保证同时只能有一个客户端对 GreatTurbo Cluster Server 10 两个节点进行配置和监控。

使用 guiadmin 的同时不能够使用其他配置工具修改集群和服务配置。

2. 配置条件

1） 必须保证在客户端和服务器端的操作系统上提供 Java 运行时环境支持，JRE1.4.2 或以上版本。
2） 在 GreatTurbo Cluster Server 10 启动以后才能使用 guiadmin 来配置和监控服务。
3） 必须保证至少一个节点已经启动 guiadmin 服务器端进程。
4） 保证服务器端的 1099 和 12345 端口不被占用；保证客户端的 1099 端口不被占用。（如果服务器端 12345 端口已经使用，那么可以修改/opt/cluster/bin/guiadminserver 脚本中"PORT"变量的值来指定新的端口号，端口号的范围为 1-65534）。
5） 如果客户端在 linux 操作系统下使用，那么要保证在 x windows 界面下运行客户端程序。
6） 需要在客户端和服务器的 PATH 中设置"java"命令的路径。
7） 检查/usr/bin 目录下的 java 命令是否正确链接到 jre 或 jdk 安装目录下的 java 命令。
8） 客户端和服务器允许 rmi 注册服务。

3. 安装、启动 guiadmin 的命令

a） Server 端

如果您购买的是正版的 GreatTurbo Cluster Server 10 产品，那么您安装 GreatTurbo Cluster Server 10 以后，guiadmin 的服务器端和客户端软件均已经自动安装。

在您使用"/etc/init.d/cluster start"命令启动 GreatTurbo Cluster Server 10 的同时 guiadmin 服务器端进程就已经启动；在使用"/etc/init.d/cluster stop"命令停止 GreatTurbo Cluster Server 10 的同时，guiadmin 服务器端进程也将被停止。

/opt/cluster/bin/guiadminserver 是启动和关闭服务器端进程的脚本，guiadminserver start 用来启动进程；guiadminserver stop 用来关闭进程。一般情况下不需要用户执行 guiadminserver 命令。如果 guiadmin 服务进程发生异常，需要用户重新启动 guiadminserver。

b） Client 端

如果用户仅仅在 GreatTurbo Cluster Server 10 的两个成员节点上运行 guiadmin 客户端软件，那么安装 GreatTurbo Cluster Server 10 即可，

不需要做额外的操作。如果用户需要远程操作，那么需要在远程客户机上（linux 操作系统）执行安装光盘里的 install_guiclient。

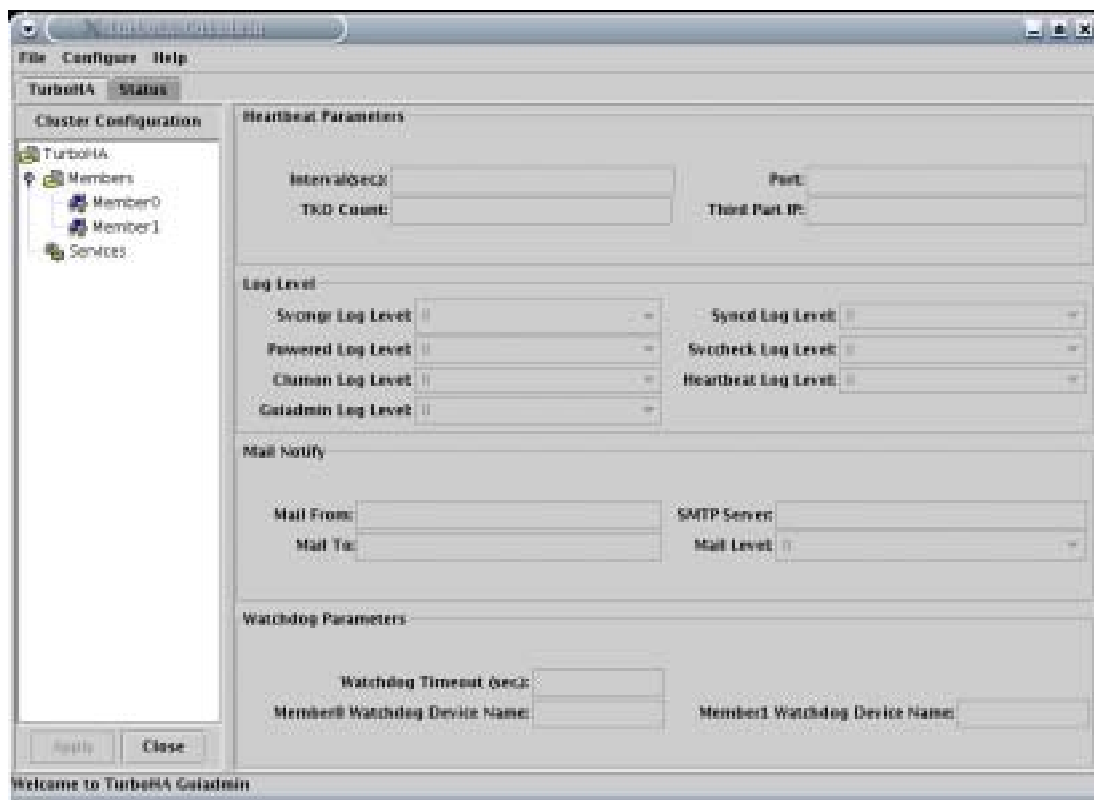
/opt/cluster/guiadminclient 是启动客户端软件的脚本。在成功安装客户端软件以后，执行此脚本文件即可启动客户端进程。

如果您需要在 windows 操作系统下使用 guiadmin 工具，那么请执行安装光盘中的/tools/guiadminclient/setup.exe。按照指示完成安装后，guiadmin 客户端程序就成功安装在 windows 操作系统下。

注意：<u>启动客户端进程以前，必须保证 Server 端进程已经被启动。</u>

4. 连接 GreatTurbo Cluster Server 10 服务器

启动客户端进程以后，将出现如下所示的图形界面：



因为没有连接到服务器，所以此时还不能够对 GreatTurbo Cluster Server 10 进行配置。用户需要和指定服务器进行连接：点击“configure”菜单中的“connect guiadmin server”项，出现如下所示的连接对话框：



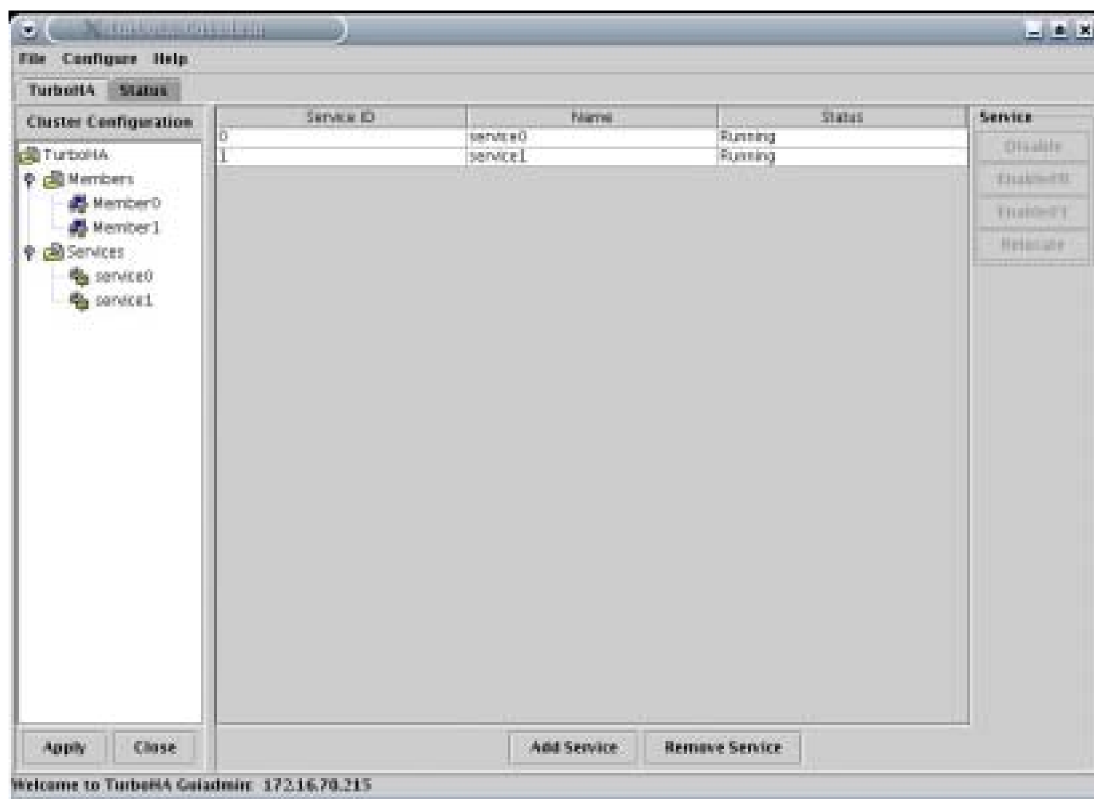
用户需要在这里填入已经启动了 guiadmin 服务进程的 GreatTurbo Cluster Server 10 两节点 IP 地址、连接的用户名和密码。正确输入所有项目以后，点击“OK”键就可以建立到服务器的连接。Guiadmin 默认先连接第一个 ip 地址，当第一个 ip 地址无法连接时尝试连接第二个 ip 地址。如果两个 ip 地址均无法连接，则返回错误信息。

为了保证安全性，guiadmin 需要客户端使用者在连接服务器端的同时输入用户名和密码。用户名和密码由管理人员使用密码生成工具设置，/opt/cluster/bin/passwdgenerate 就是专门生成 guiadmin 连接密码的工具。

运行 passwdgenerate 命令，按照提示输入用户名和密码（用户名、密码不能够为空），这样用户名和密码就保存在磁盘上。当用户连接服务器端的时候，服务器端进程会检查用户名和密码的正确性。如果用户想修改以前所设定的密码，那么可以再次运行该命令，按照提示输入用户名和密码即可。

用户在一个节点上设置完密码后，需要把/opt/cluster/lib/passwd 文件手动拷贝到另外一个节点的/opt/cluster/lib 目录下，保持两个节点的连接密码一致。

成功连接以后 GreatTurbo Cluster Server 10 两个节点的信息就全部显示出来，用户现在可以利用 guiadmin 进行配置了。下图是已经建立了连接的 guiadmin：



如果在用户使用过程中 guiadmin 服务器端进程停止服务，那么客户端程序会自动尝试连接到另外一个节点。如果连接成功，那么用户可以继续使用 guiadmin 对 GreatTurbo Cluster Server 10 进行配置；如果连接失败，guiadmin 客户端会交替尝试连接 2 个节点，直到某一节点恢复正常。

5. guiadmin 功能介绍

guiadmin 客户端大致可以分为 4 部分。即：通用信息的配置；节点信息监视；配置 HA 服务信息和节点状态监视。大部分功能的切换都在主画面左边的节点树进行，用户点击不同的节点，就会显示不同的信息。

(1)对 GreatTurbo Cluster Server 10 通用信息进行配置

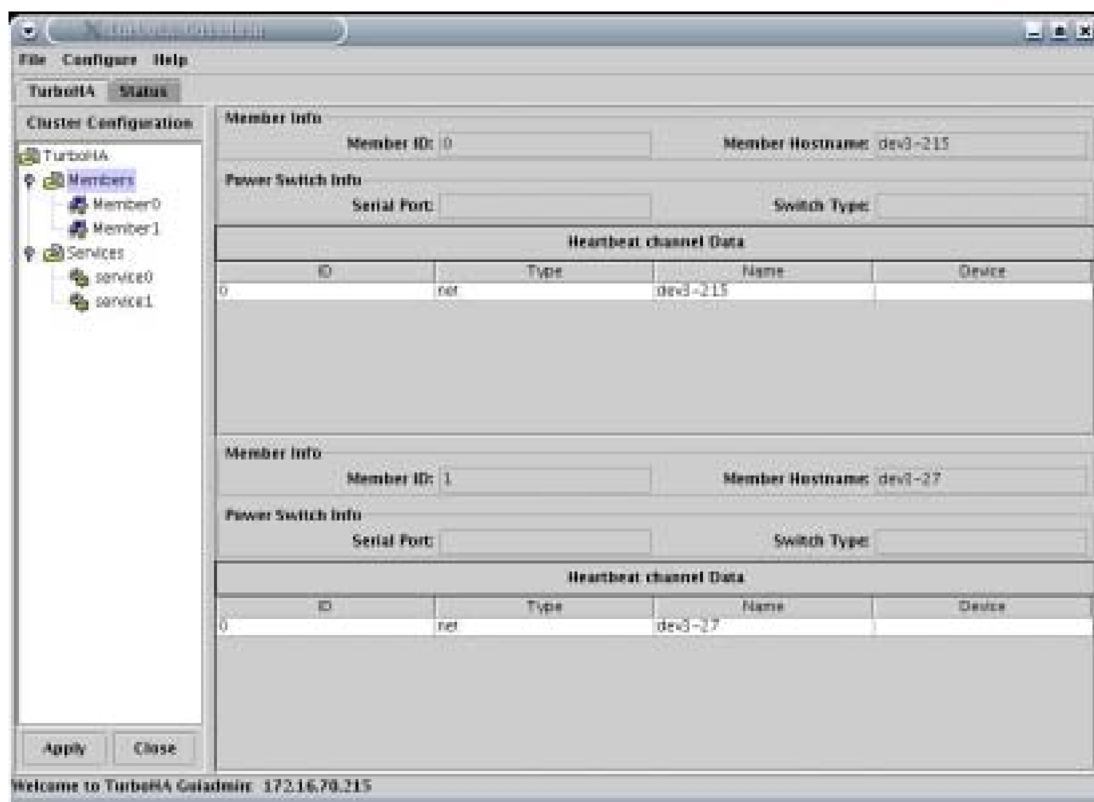
节点树中的的 GreatTurbo Cluster Server 10 项，对应的就是 GreatTurbo Cluster Server 10 通用信息。用户点击 GreatTurbo Cluster Server 10 节点，然后在右侧的界面中就可以修改信息。通用信息中的具体参数的含义请参阅第五、六、七章或者用户手册。

修改完毕以后，用户需要点击节点树下方的“Apply”键来提交给服务器端进程，这样用户的修改才能够生效。

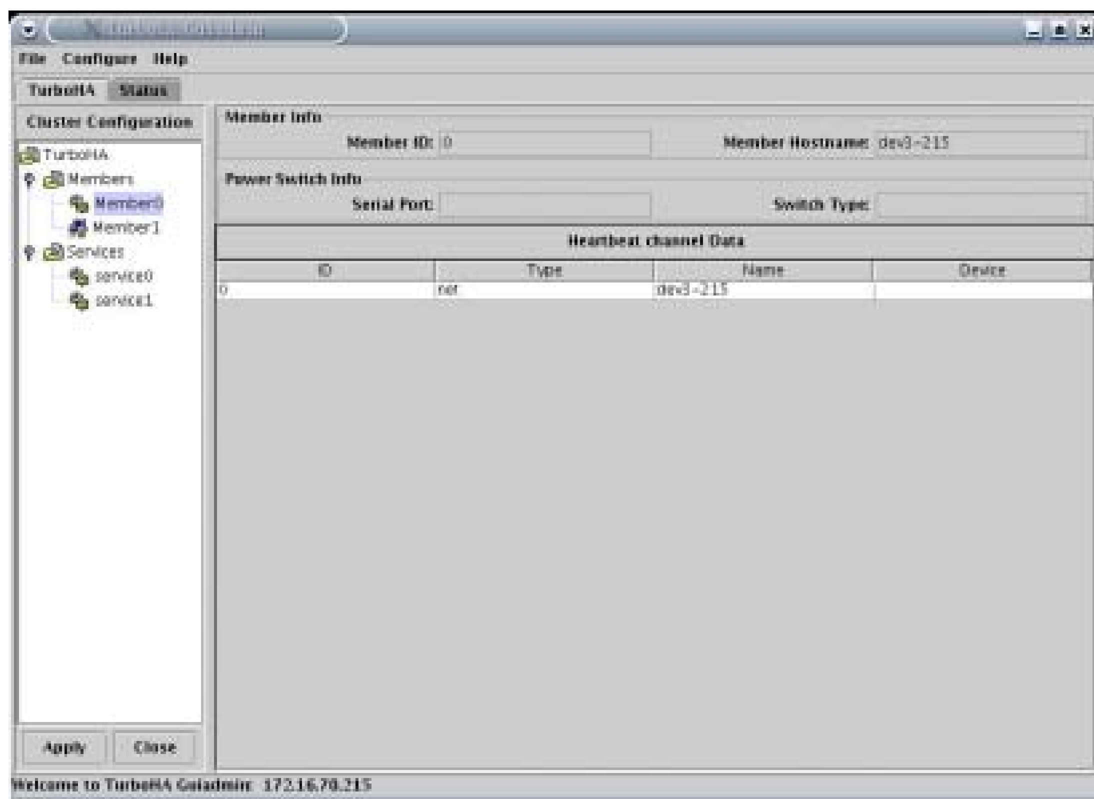
(2)节点信息监视

节点树中的 members、member0、member1 项，对应的就是节点信息。由于我们使用 member_config 里面设置了两个节点的信息，所以这里不允许用户对节点信息和心跳信息进行修改，用户仅仅可以查看信息。

图示如下：



两个节点信息显示



单一节点信息

(3)添加服务

节点树中的services项对应全部HA服务信息。在这里用户可以进行添加服务，删除服务，对服务进行动作等操作。

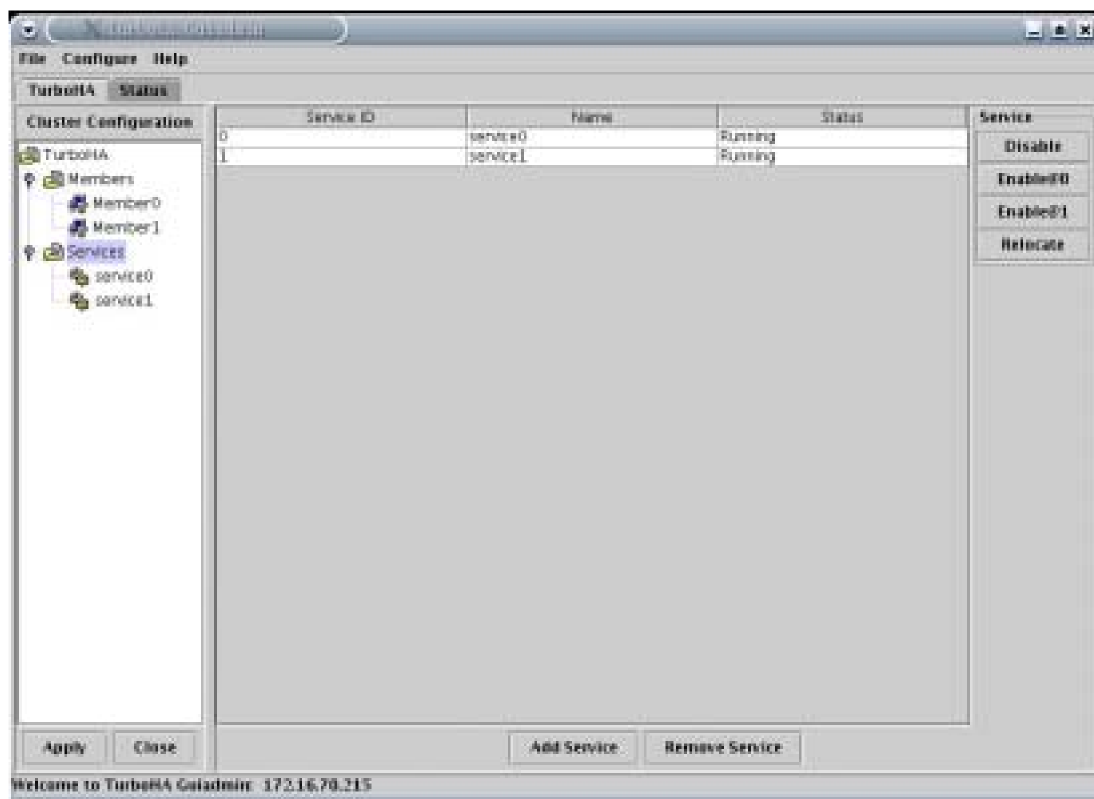
添加服务后，服务并没有配置完全。如果用户想完整地配置服务，那么需点击节点树中新增加的节点，在服务配置画面中对服务进行配置。

添加服务以后用户需要点击“Apply”键提交给服务器端进程，来使用户的操作生效。

注意：

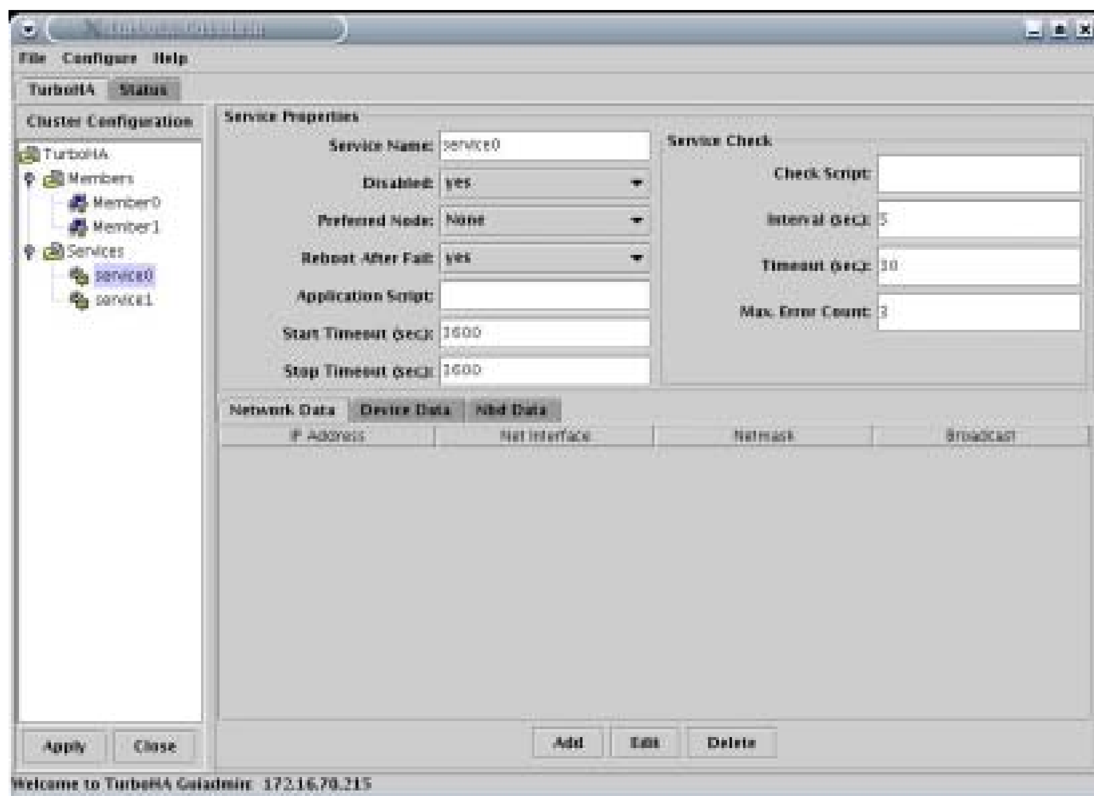
a) 一次只能够添加一个新服务。点击“Apply”键以后可以再次添加。

b) 新添加的服务名字和系统中已经存在的服务名字不能够重复。

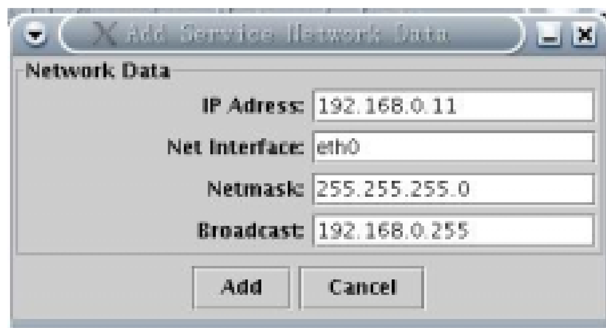


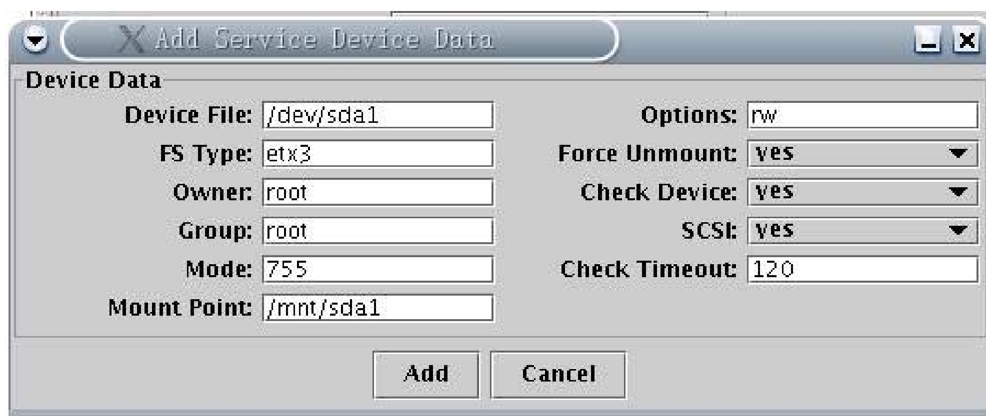
(4)服务配置信息修改

添加服务以后，在节点树中会出现新的项，节点名字对应于服务的名称。用户点击该节点，就可以在新出现的界面中修改服务信息。用户修改完配置信息以后，点击“Apply”键提交到服务器。每个参数的具体含义请参见第五、六、七、八章。



服务中可以添加、修改和删除浮动 IP，磁盘设备。点击页面下方的不同标签，即可显示不同的项目。对于 drbd 设备，我们需要用 cluadmin 来进行配置。但是在 guiadmin 中可以查看服务配置的 drbd 设备信息。







参数限制：

1) 所有服务中的浮动 IP 地址不能够重复。即 Network Data 中的 IP Address 项不能够重复。
2) 所有服务中的设备文件不能够重复，即 Device Data 中的 Device File 项不能够重复。
3) 所有服务中的设备装载点不能够重复，即 Device Data 中的 Mount Point 项不能够重复。

以上配置信息的具体含义请参见第五、六、七、八章或者用户手册。

(5)删除服务

选择处于“Not Running”状态的服务，直接点击“Remove”健就可以把服务删除。

对于“Running”状态的服务，必须先把该服务 disable，然后才能够删除。

(6)对服务的操作

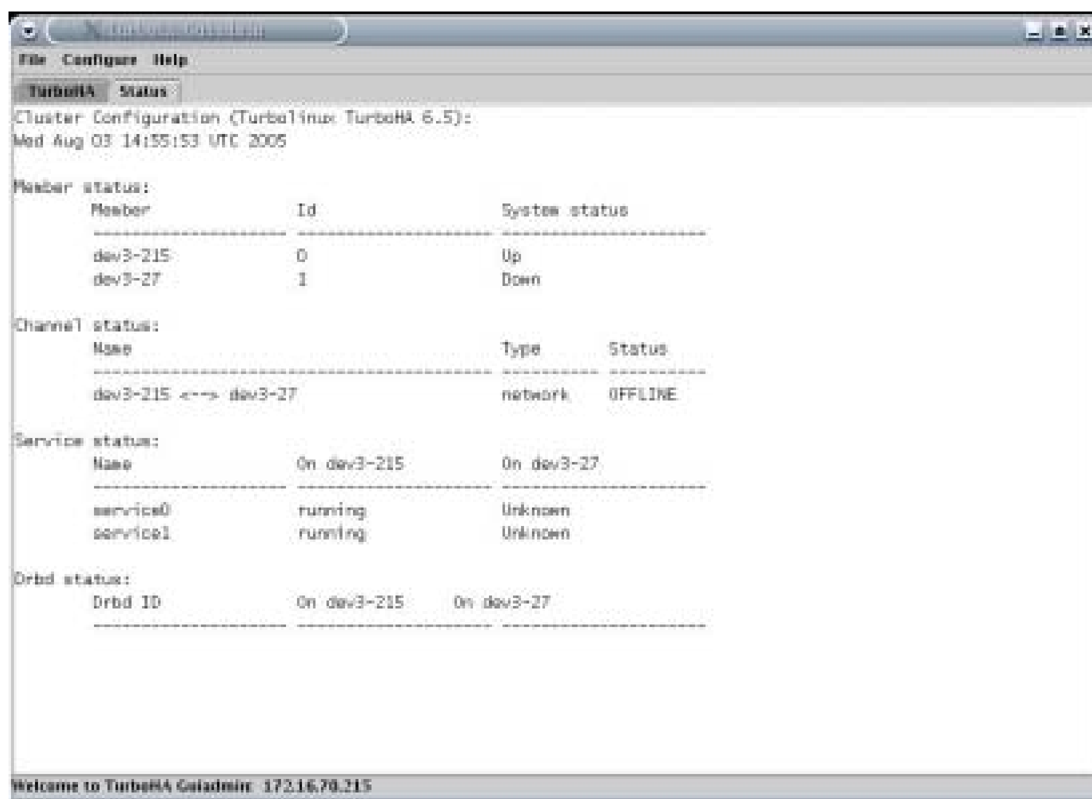
当服务处于“Running”状态时，可以点击“disable”键来使服务失效。也可以点击“Relocate”键来使服务在不同节点之间切换。

当服务处于“Not running”状态时，可以点击“Enable0”使服务在节点 0 上启动；也可以点击“Enable1”使服务在节点 1 上运行。

当服务处于“error”状态时，仅仅可以使用“disable”键来改变服务状态。

(7)GreatTurbo Cluster Server 10 两节点状态信息查询

点击客户端程序菜单下面的“Status”标签，就可以查询 GreatTurbo Cluster Server 10 两节点的状态信息。状态信息从服务器端获取，每隔 3 秒钟刷新一次。



至此， 您已经了解了 guiadmin 的大部分功能，您可以参阅第五、六、七、八章来更好的使用此工具。

6. 停止 guiadmin

(1)服务器端运行 guiadminserver stop 来停止服务器端的进程。当该进程停止以后，客户端就不能够再对 GreatTurbo Cluster Server 10 进行配置。

(2)客户端只需要点击“Close”键或者右上角的“x”来关闭客户端进程，并通知客户端释放资源。

注意：客户端进程要在服务器端停止以前关闭。

7. 卸载 guiadmin

(1)服务器端：

当您运行发行光盘上的 uninstall_cluster 以后，Guiamin 服务器端软件会随着 GreatTurbo Cluster Server 10 的卸载同时被卸载。

(2)客户端：

如果 Guiadmin 客户端软件仅仅安装在 GreatTurbo Cluster Server 10 两个节点上，那么在卸载 GreatTurbo Cluster Server 10 的同时就可以把节点上的客户端软件卸载。

如果 Guiadmin 客户端软件安装在其他 Linux 系统机器上，那么在客户机上执行“rpm –e greatturbocluster -guiadminclient”就可以把 Guiadmin 客户端软件卸载。

如果所有的服务添加完成，请到第十四章继续阅读随后内容。

第 14 章

检查 GreatTurbo Cluster Server 10 的运行状态

GreatTurbo Cluster Server 10 配置完成后，可以通过以下的方式来检测 GreatTurbo Cluster Server 10 的运行状态。

1) 检查各守护进程的运行状态。

2) 检查集群与服务的运行状态。

3) 检查 drbd 的运行状态。

1. 检查各守护进程的运行状态

可以利用以下的命令检查各守护进程的运行状态：

/etc/init.d/cluster status

GreatTurbo Cluster Server 10 启动后，正常的运行状态为：

```
[root@test1 root]# /etc/init.d/cluster status
clumon (pid 32244) is running.
hb (pid 32269) is running.
svcmgr (pid 32251) is running.
syncd (pid 32228) is running.
powerd (pid 32273) is running.
svccheck (pid 32278) is running.
```

2. 检查集群与服务的运行状态

可以利用以下的命令检查 GreatTurbo Cluster Server 10 的集群与服务的运行状态：

/opt/cluster/bin/clustat

或者在 cluadmin 中执行

cluster status 或

cluster monitor

来检查集群与服务的运行状态。

这里主要对 clustat 的使用方法进行说明，在 cluadmin 中查看状态与 clustat 基本相同。

(1) 没有配置 power switch 时的状态检查

没有配置 power switch 的情况下，GreatTurbo Cluster Server 10 正常运行时，执行 clustat 的结果为：

```
[root@test1 root]# /opt/cluster/bin/clustat
Cluster Configuration (Turbolinux GreatTurbo Cluster Server 10 6.5):
Mon Apr 11 15:05:27 CST 2005

Member status:
        Member                Id              System status
        -------------------- --------------- ---------------------
        test1                 0               Up
        test2                 1               Up

Channel status:
        Name                                  Type       Status
        ------------------------------------ ---------- ----------
        hb11 <--> hb21                        network    ONLINE
        hb12 <--> hb22                        network    ONLINE
        test1 <--> test2                      network    ONLINE
        /dev/ttyS0 <--> /dev/ttyS0            serial     ONLINE

Service status:
        Name                  On test1        On test2
        -------------------- --------------- ---------------------
        svc01                 stopped         running
```

说明：

1) 节点的状态

“Member status”表示节点的状态。

节点有以下几种状态：

Unknown: 节点状态不明

Up: 节点 GreatTurbo Cluster Server 10 已启动

Down: 节点 GreatTurbo Cluster Server 10 已停止

2) heartbeat 通道的状态

“Channel status”表示 heartbeat 通道的状态。
heartbeat 通道有以下几种状态：
Unknown: heartbeat 通道状态不明
ONLINE: heartbeat 通道正常工作
OFFLINE: heartbeat 通道断开连接

3) 服务的状态
“Service status”表示服务的状态。
服务有以下几种状态：
Unknown: 服务状态不明
starting: 服务正在启动
running: 服务正在运行
stopping: 服务正在停止
stopped: 服务已经停止
disabling: 服务正在 disable
disabled: 服务已经 disabled
not accepted: 服务网络出现问题
error: 服务状态错误

(2)配置了 power switch 时的状态检查

配置了 power switch 的情况下，GreatTurbo Cluster Server 10 正常运行时，执行 clustat 的结果为：

```
[root@test1 root]# /opt/cluster/bin/clustat
Cluster Configuration (Turbolinux GreatTurbo Cluster Server 10):
Mon Apr 11 16:13:51 CST 2005

Member status:
        Member               Id               System     Power
                                              Status     Switch
        -------------------- ---------------  ---------- ----------
        test1                0                Up         Good
        test2                1                Up         Good

Channel status:
        Name                                  Type       Status
        ------------------------------------  ---------- ----------
        hb11 <--> hb21                        network    ONLINE
        hb12 <--> hb22                        network    ONLINE
        test1 <--> test2                      network    ONLINE
```

```
        /dev/ttyS0 <--> /dev/ttyS0           serial     ONLINE

Service status:
        Name                    On test1        On test2
        -------------------- --------------- ---------------------
        svc01                   stopped         running
```

说明：

power switch 的状态:

Unknown: power switch 状态不明

Good: power switch 状态正常

Down: power switch 守护进程停止

(3)配置了磁盘镜像设备时的状态检查

当配置了磁盘镜像设备时，用带 －m 选项的 clustat 命令可以查看磁盘镜像的状态。

```
[root@test1 root]# /opt/cluster/bin/clustat －m
Cluster Configuration :
Mon Apr 11 16:28:09 CST 2005

Member status:
        Member                  Id              System status
        -------------------- --------------- ---------------------
        test1                   0               Up
        test2                   1               Up

Channel status:
        Name                                    Type       Status
        ------------------------------------ ---------- ----------
        hb11 <--> hb21                          network    ONLINE
        hb12 <--> hb22                          network    ONLINE
        test1 <--> test2                        network    ONLINE
        /dev/ttyS0 <--> /dev/ttyS0              serial     ONLINE

Service status:
        Name                    On test1        On test2
        -------------------- --------------- ---------------------
        svc01                   stopped         running

Drbd status:
```

```
Drbd ID              On test1        On test2
-------------------- --------------- ---------------------
drbd0                DP              DSS
```

说明：
“Drbd status”表示磁盘镜像的状态。
磁盘镜像有以下几种状态：
Unknown: drbd 状态不明
DP: drbd 状态为 Primary
DSS: drbd 状态为 Secondary Synced
DSU: drbd 状态为 Secondary Unsynced

3．检查 drbd 的运行状态

由于在 Linux kernel2.4 和 2.6 的环境下，drbd 的版本不同，因此，下面分别说明在这两个版本下如何检查 drbd 的运行状态。

(1) Linux kernel2.4 环境

Linux kernel2.4 的环境下，检查 drbd 的运行状态有两种方式。

1) 检查 drbd 内核模块是否在运行

可以利用以下的命令来检查 drbd 内核模块是否在运行。
/etc/init.d/drbd status
检查的结果有以下几种：
running: drbd 内核模块正常运行
stopped: drbd 内核模块已停止
error: drbd 内核模块运行错误
例如：

```
[root@test1 root]# /etc/init.d/drbd status
drbd0: running
```

2) 检查/proc/drbd

执行 cat /proc/drbd 可以查看 drbd 的更详细的信息。

例如：

```
[root@test1 root]# cat /proc/drbd
version: 0.6.8 (api:63/proto:62)

0:  cs:Connected  st:Primary/Secondary  ns:104397  nr:0  dw:9
dr:104521 pe:0 ua:0
```

说明：

Field	说明	值
Cs	连接状态	o Connected：一切正常 o Unconfigured：过渡状态，设备在等待配置(insmod 后的状态) o StandAlone：过渡状态，配置完 disk 没配 net。 o Unconnected：连接模块时的过渡状态，或连接失败时的状态 o Timeout：发送数据包 timeout o BrokenPipe：发送数据包错误 o WFConnection：连接之前的状态 o WFReportParams：发送参数给对方 o SyncPaused：等待高优先级的 group 先 sync 完 o SyncingAll：正将主节点的所有模块复制到次级节点上 o SyncingQuick：通过复制已被更新的模块（在次级节点短暂离开集群的情况下）来更新次级节点
St	状态	可能的值为： o Primary/Secondary o Primary/Unknown o Secondary/Secondary o Secondary/Unknown
Ns	发送的数据	KB
Nr	接收的数据	KB
Dw	磁盘写入的数据	KB
Dr	磁盘读取的数据	KB
Of	运行中(过时)的请求	请求数
Pe	待解决的请求	请求数
Ua	未答复的请求	请求数

(2) Linux kernel2.6 环境

在 Linux kernel2.6 的环境下，可以用以下的命令查看 drbd 的状态。

```
[root@test1 root]# /etc/init.d/drbd status
drbd driver loaded OK; device status:
version: 0.7.11 (api:77/proto:74)
```

```
SVN Revision: 1912M build by root@qa3-127, 2005-08-25 10:20:28
 0: cs:Connected st:Primary/Secondary ld:Consistent
    ns:20387 nr:366908 dw:387295 dr:14614 al:47 bm:139 lo:0
pe:0 ua:0 ap:0
```

说明：

Field	说明	值
Cs	连接状态	o Connected：一切正常 o Unconfigured：过渡状态，设备在等待配置(insmod 后的状态) o StandAlone：过渡状态，配置完 disk 没配 net。 o Unconnected：连接模块时的过渡状态，或连接失败时的状态 o Timeout：发送数据包 timeout o BrokenPipe：发送数据包错误 o NetworkFailure：网络错误 o WFConnection：连接之前的状态 o WFReportParams：发送参数给对方 o SkippedSyncS：应当为同步的源端，但用户选择不同步 o SkippedSyncT：应当为同步的目的端，但用户选择不同步 o WFBitMapS：过渡状态，同步源端同步前等待 bitmap o WFBitMapT：过渡状态，同步目的端同步前等待 bitmap o SyncSource：同步的源端 o SyncTarget：同步的目的端 o PausedSyncS：同步源端暂停 o PausedSyncT：同步目的端暂停
St	状态	可能的值为下列值的组合： o Primary o Secondary o Unknown
ld	物理磁盘的状态	o Consistent：数据一致 o Inconsistent：数据不一致
Ns	发送的数据	KB
Nr	接收的数据	KB
Dw	磁盘写入的数据	KB
Dr	磁盘读取的数据	KB
Al	访问 log 的数目	个数
Bm	写 bitmap 的数目	个数
Lo	等待本地磁盘来	请求数

	标记操作完成的请求	
pe	Pending(等待应答的请求)	请求数
ua	没有应答的请求	请求数
ap	等待结束的请求	请求数

4. 检查 ipvs 调度情况

可以使用 clustat 命令来观察 ipvs 的调度情况：

1) 使用“/opt/cluster/bin/clustat –L”来列出 LVS 信息。

2) 使用“/opt/cluster/bin/clustat –LS”来列出统计数据

3) 使用“/opt/cluster/bin/clustat –LR”来列出速率信息

第 15 章

卸载GreatTurbo Cluster Server 10

卸载GreatTurbo Cluster Server 10的过程如下：

1) 卸载GreatTurbo Cluster Server 10
2) 卸载drbd
3) 卸载GreatTurbo Cluster Server10 realserver包
4) 卸载guiadmin客户端

1. 卸载GreatTurbo Cluster Server 10

您可以利用GreatTurbo Cluster Server 10提供的uninstall_cluster工具来卸载GreatTurbo Cluster Server 10。uninstall_cluster安装在/sbin目录下。

注意：

卸载的操作请在两个节点上分别执行。

卸载的方法如下：

1) 卸载之前请先停止GreatTurbo Cluster Server 10。GreatTurbo Cluster Server 10停止的方法如下：

```
/etc/init.d/cluster stop
```

例如：

```
[root@test1 root]# /etc/init.d/cluster stop
--------- Performing cluster stop ---------------------
---- Performing guiadmin server stop ----
Stopping guiadmin server: done.
---- Completed guiadmin server stop ----
Shutting down clumon:done
Sending stop message to svcmgr: done.
Waiting for Cluster Daemons to exit.
Stopping syncd: done.
Stopping hb: done.
```

```
Stopping svcmgr: done.
Stopping powerd: done.
Stopping svccheck: done.
____ Performing drbd stop ____
drbd: 'drbd0' already Secondary
drbd: module has been unloaded
____ Completed drbd stop ____
-------- Completed cluster stop --------------------
```

2) 执行卸载工具 uninstall_cluster 来卸载 GreatTurbo Cluster Server 10。如果您安装了 drbd，请在提示是否卸载 drbd 时选择 y。

```
[root@test1 root]# ./uninstall_cluster
Following RPMs will be uninstalled:
  *) GreatTurbo Cluster Server
  *) GreatTurbo Cluster Server Agents
  *) drbd(if installed)
Do you want to continue? (y/n) [y]: y
Do you want to uninstall drbd? (y/n) [y]: y
GreatTurbo Cluster Server uninstallation finished.
```

2. 卸载 drbd

如果您安装了 drbd，执行 uninstall_cluster 会卸载 drbd 软件包。如果您要单独卸载 drbd，请进行以下操作：

(1)Linux kernel2.4 环境

在 Linux kernel2.4 环境下，卸载 drbd 软件包，请执行以下的命令：

```
[root@test1 root]# rpm -e drbd
```

(2)Linux kernel2.6 环境

在 Linux kernel2.6 环境下，卸载 drbd 软件包，请执行以下的命令：

其中，xxx 代表 Linux kernel 版本号。

```
[root@test1 root]# rpm -e drbd-km-xxx
[root@test1 root]# rpm -e drbd
```

3. 卸载 GreatTurbo Cluster Server10 realserver 包

卸载的方法如下：

1) 卸载之前请先恢复 realserver 配置,根据选择的负载技术执行不同的命令：

```
[root@test1 root]# /etc/init.d/lbrealserver-dr stop
```

2) 执行卸载工具 uninstall_realserver 来卸载 GreatTurbo Load Balance Server-realserver

```
[root@test1 root]# uninstall_realserver
```

4. 卸载 guiadmin 客户端

执行 install_cluster 在 GreatTurbo Cluster Server 10 集群的两个节点上安装的 guiadmin 客户端，在执行 uninstall_cluster 时会自动卸载。

如果在别的机器上执行 install_guiclient，单独安装了 guiadmin 客户端，请执行以下的命令进行卸载：

```
rpm -e greatturbocluster-guiadminclient
```

附录

GreatTurbo Cluster Server 10 的维护

如果您还有什么疑问或者是您认为 GreatTurbo Cluster Server 10 在运行中出现问题，可以首先对照 FAQ 进行自行处理。如果发现仍然不能处理相应问题。可以与拓林思软件有限公司进行联系。

1. GreatTurbo Cluster Server 10 的日志信息

GreatTurbo Cluster Server 10 的日志信息存放在两个节点的/var/log/cluster 文件中，如果您的系统出现问题，您可以首先阅读两个节点的/var/log/cluster 日志文件，通过日志可以初步定位一些基本的故障和原因。

注意：

1) 在 GreatTurbo Cluster Server 10 运行期间，请不要手动修改和删除两个节点上的/var/log/clluster 文件，以免 log 信息丢失。
2) 如果您发现/var/log/cluster 日志文件较大,可以在两个节点上分别直接运行“/opt/cluster/bin/clusterclear” 命令来备份并清空日志文件。
3) 如果您发现/var/log/cluster 日志文件没有正常记录日志信息，可以执行“/etc/init.d/syslog restart”重新激活/var/log/cluster 文件的日志功能。

2. FAQ

1) Q: 如何查看 GreatTurbo Cluster Server 10 安装的是哪一个发行版本？
   A: 可以使用 rpm -qa |grep greatturbocluster 命令来查看具体的版本号。

2) Q: 如果配置了第三方 IP 地址，但是第三方 IP 地址由于某种原因发生了一段时间（比如说 1 天）的故障，在这端时间内是否会影响 GreatTurbo Cluster Server 10 的正常使用？
   A: 在这段时间内，如果没有节点发生故障，则不会影响 GreatTurbo Cluster Server 10 的正常使用，否则由于一个节点的故障（例如节点掉电或死机等），可能会导致另一个节点的反复 reboot 现象，从而会导致用户的业务受到影响。所以，如果不能保证第三方参考 IP 永久持续有效，则可以不配置第三方 IP。

3) Q: 如果不配置第三方 IP 地址，会对 GreatTurbo Cluster Server 10 系统有影响吗？

A: 如果两个节点之间配置了三条以上的心跳通道，那么这三条心跳通道全部故障的几率是非常小的。在这个前提下，如果用户实在是不具备永久持续有效的第三方参考 IP，则可以不配置第三方 IP 地址，在使用 member_config 工具的配置过程中将第三方 IP 配置选项选择“no”。如果用户具备永久持续有效的第三方参考 IP 时，则最好还是配上，在 member_config 的配置过程中将第三方 IP 配置选项选择“yes”，并输入地址值。

4) Q: 什么是永久持续有效的第三方参考 IP？
A：a）该永久持续有效，不会 down 掉，也不会短时故障。
b）从 GreatTurbo Cluster Server 10 的节点能够 ping 通该 IP。

5) Q: GreatTurbo Cluster Server 10 为什么要配置多个心跳通道？
A: GreatTurbo Cluster Server 10 配置三条以上的网络通道是出于最大可能提高系统可用性的考虑。通过三条以上的通道来保证硬件上的更高冗余性。GreatTurbo Cluster Server 10 使用通道来同步心跳信息和发送命令。如果第一条通道正常，则 GreatTurbo Cluster Server 10 会选择第一条通道进行通讯，如果第一条通道故障，那么 GreatTurbo Cluster Server 10 会选择下一条正确的通道进行通讯，以此类推。也就是说 TuboGreatTurbo Cluster Server 10 按照 member_config 工具配置通道时的顺序来使用通道。

6) Q: GreatTurbo Cluster Server 10 除了将直连网线和串口配置成通道之外，为什么还需要将服务使用的外网网卡也配置成通道？
A：这是为了进一步保证服务的高可用性，以避免当所有的直连网线和串口都发生故障时会发生的裂脑现象，以保证应用数据的一致性。
需要强调的是，将服务使用的外网网卡配置成通道并不会影响服务的网络带宽，因为 GreatTurbo Cluster Server 10 会自动优先使用直连网线通道。只有所有的直连网线都发生故障时，GreatTurbo Cluster Server 10 才会使用外网网卡。

7) Q: 如何升级 GreatTurbo Cluster Server 10？
A: 如果由于某种原因，用户需要升级 GreatTurbo Cluster Server 10 到更高发行版本时，只需先按照手册卸载旧版本的 GreatTurbo Cluster Server 10，然后按照手册重新安装新版本的 GreatTurbo Cluster Server 10 即可。基本顺序是：1.卸载旧版本；2.安装新版本；3.如果原先配置的应用不需要更改或者也不需要增加新的应用，则无需重新配置，直接可以启动 GreatTurbo Cluster Server 10 进行使用。

8) Q: 如果用户应用的 CPU 负载较重时,需要调整 GreatTurbo Cluster Server 10 的参数吗？
A: 如果用户应用的 CPU 负载较重,且持续在 n*5 秒钟以内都维持在 100%以上时，需要使用 cluadmin 工具来进行调整 GreatTurbo Cluster Server 10 同步心跳的参数。进入 cluadmin，输入 cluster heartbeat interval n 将时钟同步的间隔时间设置成 n 秒钟。

如果用户的应用的 CPU 负载不会达到 100%，或者是 100%CPU 负载持续时间在 15（3*5）秒以内，则不需要改动参数。

9) Q: GreatTurbo Cluster Server 10 是否能够检测网卡的硬件连接情况？
A: 能够。GreatTurbo Cluster Server 10 通过配置，能够检测服务所使用的网卡是否正常连接到交换机或者 HUB 上。
注意：待检测网卡的网线另一段必须是连接在交换机或者 HUB 上。

10) Q: GreatTurbo Cluster Server 10 服务迁移时间是否可以进行调整？
A: 可以。通过调整服务检测到错误的时间来调整服务迁移时间。由于服务检测到错误的时间＝服务检测的间隔时间 * 服务连续出错的次数，所以在配置服务的检测脚本时，通过调整 check interval 以及 check count 参数，可以控制服务检测到错误的时间。

通常，习惯于将 check count（单位为次）设置为 3，表示连续 3 次出错后才确认服务的确出错，以保证检测的可靠性。而 check interval（单位为秒）的设置由用户的应用决定。

一定要根据用户的应用要求来合理的设置 check interval 的值。

如果应用更着重于稳定性/可靠性（属于电信/银行等关键业务的应用或应用本身的负载较重等），建议将 interval 的值适当调大，如设置 check interval 在 10 秒-30 秒之间，当服务出错，服务最少将在 30 秒-90 秒之后进行服务的迁移。

如果应用更着重于性能（应用负载较轻或者应用要求极短的服务不可用时间），建议将 interval 的值适当调小，如设置 check interval 在 3 秒-10 秒之间，当服务出错，服务最少将在 9 秒-30 秒之后进行服务的迁移。

对于通常的应用，一般将 check interval 设置成 5 秒，当服务出错，服务最少将在 15 秒之后进行服务的迁移。

11) Q:当服务出现 error 状态时该如何处理？
A: 当服务启动或停止失败时，服务将处在错误状态，处在错误状态下的服务，其资源的状态将无法确定（例如，它的部分资源已经释放，但某些服务资源仍然将在所有者节点上进行配置）。GreatTurbo Cluster Server 10 中，以下的情形将导致服务的状态为 error：①启用服务失败；②停止服务失败(多为服务检测失败时停止服务)；③禁用服务失败；④切换服务失败；⑤停止 GreatTurbo Cluster Server 10 失败(即/etc/init.d/cluster stop 失败)。

由于处理错误状态下的服务可能有资源没有释放，所以处理错误状态中的服务必须十分小心。如果服务资源仍然在所有者节点上配置，那么在另一个集群节点上启动服务可能会造成严重的问题。例如，如果某一文件系统已经加载到它的所有者节点上，您又在另一个集群节点上启动该服务，则该文件系统将可能同时在两个系统上被加载，这样可能导致数据损坏。所以，您必须在确认所有的服务资源都已经完全释放掉了之后才能在另一个集群节点上重新启用服务。

对于 errror 状态服务，若按服务配置的重启策略自动重启后仍然不成功，就需要手动干预。手动干预处理错误状态的服务的过程一般如下：

①定位错误的原因。当服务的状态变为 error 时，首先要定位错误的原因，要从 GreatTurbo Cluster Server 10 的日志中分析。一般服务启动和停止的过程分为几步(例如，启动服务时会有 mount 分区，启动活动 IP，执行启动脚本等几步)，每一步出错都会打印日志，要从日志中分析具体是哪一步出错了。
②排除故障。当找到错误的原因后，需要排除故障。一般是通过先 modify，使服务配置正确，然后 disable 操作，使服务能够成为 disabled 状态，若操作成功就可以进行 enable 或 delete 等其它的操作了。在 modify 的过程中，建议修改服务配置使 Reboot After Fail 为 no，以避免在 modify 服务仍有问题的时候进行无谓的重启。例如，如果是启动/停止脚本写的有错误，则需要纠正错误并测试正确等。
③释放资源，恢复系统。故障排除后，需要释放所有的没有释放的服务资源，并重新恢复系统。

12) Q: 使用磁盘镜像设备(drbd)时，如何配置 Oracle 的服务？
A: 当使用磁盘镜像设备时，如果想配置 Oracle 的服务，建议将 Oracle 在 GreatTurbo Cluster Server 10 的两个节点上分别安装，即不安装在磁盘镜像设备上，而安装在节点自己的磁盘设备上，安装时先不要建库，在安装完后再单独建库。将 Oracle 的库文件建在磁盘镜像设备上，建库时，先在一个节点上 mount 磁盘镜像设备、建库，完成后，再在另一个节点上 mount 磁盘镜像设备、以和前一个节点上完全相同的操作建库，并选择覆盖原来的库。这样，Oracle 就在磁盘镜像设备上安装完成了，然后先手动试一下能否启动 Oracle，如果手动能够启动 Oracle，就可以在 GreatTurbo Cluster Server 10 中配置基于磁盘镜像设备的 Oracle 服务了。

3. Log 收集工具的使用方法

当 GreatTurbo Cluster Server 10 发生故障时，可以使用 GreatTurbo Cluster Server 10 附带的 log 收集工具 clulogc 收集 log 和系统的信息，将收集到的信息发送给拓林思公司，以便更有效地定位故障的原因。

Log 收集工具的存放位置：/opt/cluster/bin/clulogc

Log 收集工具的使用方法：/opt/cluster/bin/clulogc [OPTIONS]

命令选项	-a	收集所有的信息
	-t	收集 GreatTurbo Cluster Server 10 相关的信息
	-s	收集 syslog
	-c	收集 core 文件
	-o	收集系统的信息
	-h	显示帮助信息
参数选项	-n clulog_rotate_num	指定收集 GreatTurbo Cluster Server 10 log 的文件个数
	-p syslog_path	指定 syslog 的路径
	-r syslog_rotate_num	指定收集 syslog 的文件个数

至少要指定一个命令行选项，参数选项可以指定也可以不指定。

当不指定任何参数选项时，表示收集所有的 GreatTurbo Cluster Server 10 的 log 和 syslog，并且 syslog 的路径为/var/log/message。

如果 syslog 的存放位置不在/var/log/message，需要利用-p 选项指定 syslog 的完整的路径。

假设要收集所有的 log 信息,并且只收集当前的 GreatTurbo Cluster Server 10 的 log 和 syslog，命令行应为：

/opt/cluster/bin/clulogc －a －n 1 －r 1

假设只收集 GreatTurbo Cluster Server 10 相关的所有信息，命令行应为：

/opt/cluster/bin/clulogc -t

收集到的 log 信息的存放位置：/opt/cluster/log/

收集到的 log 信息的文件名格式：机器名-年-月-日.tar.gz

例如机器名为 test1，收集日期为 2005 年 1 月 1 日，收集到的 log 信息的文件名为：test1-2005-01-01.tar.gz

注意：

1) log 收集工具需要在两个节点上分别执行。
2) 为了收集比较全面的信息，请用/opt/cluster/bin/clulogc －a 来收集信息；如果用-a 选项收集到的 log 文件过大，请分别用-t, -s, -c, -o 选项来收集信息。

4. 联系拓林思软件有限公司

1) 邮件联系：support@turbolinux.com.cn。

   请在邮件中详细描述 GreatTurbo Cluster Server 10 的版本信息，故障现象，并在附件中附上用 log 收集工具收集的 log 信息。

2) 电话联系：01065054020